



UNIVERSIDAD DE CUENCA

Facultad de Ciencias Económicas y Administrativas

Carrera de Economía

“Utilización de Redes Neuronales Artificiales (RNA) como alternativa para el pronóstico del Índice de Precios al Consumidor (IPC) y contraste de resultados con modelos estadísticos tradicionales (ARIMA y VAR) para el Ecuador periodo 2000-2017.”

Trabajo de titulación previo a la obtención del

Título de Economista.

Modalidad: Artículo Académico

AUTOR:

Anthony Daniel Mora Saavedra

CI. 0105532782

DIRECTOR:

Eco. Carlos Julio Rivera Bautista

C.I. 0102408234

Cuenca – Ecuador

2018

RESUMEN

En la presente investigación se evalúa la calidad de pronóstico ofrecida por tres modelos diferentes, dos de ellos de uso tradicional en las ciencias económicas (modelo ARIMA y VAR) y un modelo de redes neuronales artificiales (RNA) para la serie del índice de precios al consumidor (IPC) del Ecuador para el periodo 2000-2017. Los modelos que mejores resultados ofrecen dentro de cada familia son el modelo SARIMA (1, 1,1), modelo VAR con un solo rezago y el modelo RNA con 1 capa oculta, 8 neuronas ocultas y función de activación lineal en cada capa. Como principales hallazgos, se destaca la gran eficacia de pronóstico que el modelo SARIMA (1, 1,1) tiene, tanto para el corto como para el mediano plazo, también se debe destacar el aporte que el modelo de RNA hace a la calidad de pronóstico multivariante cuando es utilizado en conjunto con el modelo VAR, mejorando la raíz del error cuadrático medio. Para analizar la calidad de pronóstico de los diferentes modelos, a más de los indicadores estadísticos tradicionales (error cuadrático medio y su raíz, error porcentual medio, promedio error porcentual absoluto y su desviación y U1 de Theil) también se realiza un pronóstico recurrente, el cual muestra que todos los modelos con los que se ha realizado el pronóstico se muestran estables.

Palabras clave: pronóstico, pronóstico recurrente, redes neuronales, algoritmos, neuronas, vectores autoregresivos, estacionariedad, integración, autocorrelación serial.

Código JEL: C45, C22, C14, E37,

ABSTRACT

In the present investigation the quality of forecast offered by three different models is evaluated. Two of them of traditional use in economic sciences (model ARIMA and VAR) and a model of artificial neural networks (ANN) for the series of the consumer price index (CPI) of Ecuador for the period 2000-2017. The models that offer the best results within each family are the SARIMA model (1,1,1), VAR model with a single lag and the ANN model with 1 hidden layer, 8 hidden neurons and linear activation function in each layer. As the main findings, it is highlighted the great prognosis efficiency that the SARIMA model (1,1,1) has, both for the short and medium term. It is also important to highlight the contribution that the ANN model makes to the quality of multivariate prognosis when used together with the VAR model, improving the root mean quadratic error. To analyze the prognosis quality of the different models, in addition to the traditional statistical indicators (mean quadratic error and its root, mean percentage error, average absolute percentage error and its deviation and Theil U1) a recurrent prognostication is also made. This prognostication shows that all the models which the prognosis has been made with are stable.

Keywords: forecast, recurrent forecast, neural networks, algorithms neurons, autoregressive vector, mobile average, stationarity, cointegration, serial autocorrelation.

JEL Codes: C45, C22, C14, E37,



Índice de contenido

1. INTRODUCCIÓN.....	9
1.1. Revisión de la literatura	12
1.2. Metodología.....	18
1.2.1. Modelo autoregresivo integrado de media móvil (ARIMA).	19
1.2.2. Modelo de vectores autoregresivos (VAR).....	21
1.2.3. Modelo de redes neuronales artificiales (RNA)	22
1.2.4. Pronóstico	29
1.3. Datos y tratamiento.....	30
2. RESULTADOS Y DISCUSIÓN.....	33
3. CONCLUSIONES Y RECOMENDACIONES.....	44
4. BIBLIOGRAFÍA	49
5. ANEXOS	52

Índice de Gráficos

Gráfico 1: Metodología Box – Jenkins para la selección del esquema ARIMA (p, d, q)20	
Gráfico 2: Simplificación matemática de una neurona (neurona artificial)	23
Gráfico 3: Topología perceptrón multicapa	26
Gráfico 4: Funcionamiento del algoritmo back-propagation	27
Gráfico 5: Red neuronal artificial recurrente.	29
Gráfico 6: Pronostico de modelos SARIMA mediano y corto plazo	34
Gráfico 7: Pronostico de modelos VAR mediano y corto plazo	38
Gráfico 8: Pronostico modelo RNA corto y mediano plazo.....	42
Gráfico 9: RECM diferentes modelos en cada iteración del pronóstico recurrente.	44

Índice de Tablas

Tabla 1: Resultados de los modelos ARIMA.	34
Tabla 2: Resultados de los modelos VAR	37
Tabla 3: Resultado modelo RNA de corto y mediano plazo	41
Tabla 4: Resultados del pronóstico de los diferentes modelos.....	43

Índice de Gráficos Anexos

Gráfico Anexo 1: IPC 2000-2017	53
Gráfico Anexo 2: IPC 2000-2003	53
Gráfico Anexo 3: IPC 2004-2017	54
Gráfico Anexo 4: Variables explicativas.....	55
Gráfico Anexo 5: Variable en diferencia	55
Gráfico Anexo 6: Variables en diferencia normalizado agrupadas.....	56



Gráfico Anexo 7: Variables en diferencia normalizadas.....	56
Gráfico Anexo 8: Variables en diferencia acotada entre 0 y 1 agrupadas	57
Gráfico Anexo 9: Variables en diferencia acotada entre 0 y 1	57
Gráfico Anexo 10: Variables en diferencia acotada entre -1 y 1 agrupadas.....	57
Gráfico Anexo 11: Variables en diferencia acotada entre -1 y 1.....	58
Gráfico Anexo 12: Prueba KMO y esfericidad de Bartlett.....	58
Gráfico Anexo 13: Varianza explicada por cada componente y grafico de sedimentación.....	59
Gráfico Anexo 14: Matriz de componentes	59
Gráfico Anexo 15: Matriz de componentes rotados.....	60
Gráfico Anexo 16: PAC y FAC Modelo ARIMA (1, 1,1).....	61
Gráfico Anexo 17: Histograma de errores modelo ARIMA (1, 1,1)	62
Gráfico Anexo 18: Estimación modelo SARIMA (2, 1,1)	62
Gráfico Anexo 19: Histograma de errores modelo SARIMA (2, 1,1).....	63
Gráfico Anexo 20: PAC y FAC Modelo SARIMA (2, 1,1)	63
Gráfico Anexo 21: Estimación modelo SARIMA (1, 1,1) mediano plazo	64
Gráfico Anexo 22: Histograma de errores modelo SARIMA (1, 1,1) mediano plazo.....	64
Gráfico Anexo 23: PAC y FAC Modelo SARIMA (1, 1,1) mediano plazo	65
Gráfico Anexo 24: Estimación modelo SARIMA (1, 1, 1) corto plazo	66
Gráfico Anexo 25: Histograma de errores modelo SARIMA (1, 1,1) corto plazo	66
Gráfico Anexo 26: PAC y FAC Modelo SARIMA (1, 1,1) corto plazo	67
Gráfico Anexo 27: Resultados VAR exploratorio.....	68
Gráfico Anexo 28: Test raíz unitaria y autocorrelación serial	68
Gráfico Anexo 29: Resultados VAR variables de la revisión de la literatura 1 rezago..	69
Gráfico Anexo 30: Test raíz unitaria y autocorrelación serial 1 rezago.	70
Gráfico Anexo 31: Especificación VAR variables de la revisión de la literatura 14 rezagos.	70
Gráfico Anexo 32: Resultados VAR variables de la revisión de la literatura 14 rezagos.	70
Gráfico Anexo 33: Test raíz unitaria y autocorrelación serial 14 rezagos.	71
Gráfico Anexo 34: Resultados VAR componentes principales.	71
Gráfico Anexo 35: Test raíz unitaria y autocorrelación serial	72
Gráfico Anexo 36: Perceptrón multicapa y red recurrente.....	73
Gráfico Anexo 37: Pronósticos modelos RNA perceptrón multicapa y red recurrente.	73
Gráfico Anexo 38: Pronósticos modelos RNA perceptrón multicapa y red recurrente.	74
Gráfico Anexo 39: Función sigmoidea logarítmica y tangente hiperbólica.	75
Gráfico Anexo 40: Pronósticos modelos RAN sigmoidea logarítmica y tangente hiperbólica.	76
Gráfico Anexo 41: Arquitectura con 2 capas ocultas.....	76
Gráfico Anexo 42: Pronósticos modelos RNA dos capas ocultas.	77
Gráfico Anexo 43: Reducción del número de neuronas ocultas.....	78
Gráfico Anexo 44: Resultados modelos RNA diferente número de neuronas.	79
Gráfico Anexo 45: Pronostico modelos RNA diferente número de neuronas.....	79
Gráfico Anexo 46: Modelo RNA componentes principales.....	80
Gráfico Anexo 47: Pronostico modelo RNA componentes principales.....	80



Índice de Tablas Anexas

Tabla Anexo 1: Medidas estadísticas de calidad de pronóstico	52
Tabla Anexo 2: Glosario de variables	54
Tabla Anexo 3: Especificación VAR exploratorio	68
Tabla Anexo 4: Especificación VAR variables de la revisión de la literatura 1 rezago .	69
Tabla Anexo 5: Especificación VAR componentes principales.....	71
Tabla Anexo 6: Descripción modelos RNA.	72
Tabla Anexo 7: Resultados modelos RNA sigmoidea logarítmica y tangente hiperbólica.	75
Tabla Anexo 8: Resultados modelos RNA dos capas ocultas.	77
Tabla Anexo 9: Resultados modelo RNA componentes principales.....	80



Cláusula de Licencia y Autorización para Publicación en el Repositorio Institucional

Yo, Anthony Daniel Mora Saavedra en calidad de autor y titular de los derechos morales y patrimoniales del trabajo de titulación "Utilización de Redes Neuronales Artificiales (RNA) como alternativa para el pronóstico del Índice de Precios al Consumidor (IPC) y contraste de resultados con modelos estadísticos tradicionales (ARIMA y VAR) para el Ecuador periodo 2000-2017." reconozco a favor de la Universidad de Cuenca una licencia gratuita, intransferible y no exclusiva para el uso no comercial de la obra, con fines estrictamente académicos.

Asimismo, autorizo a la Universidad de Cuenca para que realice la publicación de este trabajo de titulación en el Repositorio Institucional, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Cuenca, Diciembre del 2018

Anthony Daniel Mora Saavedra

CI. 0105532782



Cláusula de Propiedad Intelectual

Yo, Anthony Daniel Mora Saavedra, autor del trabajo del Trabajo de Titulación "Utilización de Redes Neuronales Artificiales (RNA) como alternativa para el pronóstico del Índice de Precios al Consumidor (IPC) y contraste de resultados con modelos estadísticos tradicionales (ARIMA y VAR) para el Ecuador periodo 2000-2017", certifico que todas las ideas, opiniones y contenidos expuestos en la presente investigación son de exclusiva responsabilidad de su autor.

Cuenca, Diciembre del 2018

Anthony Daniel Mora Saavedra

CI. 0105532782

1. INTRODUCCIÓN

El Índice de Precios al Consumidor (IPC) es una de las variables macroeconómicas más importantes para los agentes que interactúan dentro de una economía (hogares, empresas, estado), ya que este es un reflejo de la variación que presentan los precios de los diferentes bienes y servicios que se ofertan en el mercado a través del tiempo (Naciones Unidas, 2017). Debido a la importancia de este indicador, el Instituto Nacional de Estadísticas y Censos (INEC) lleva a cabo la actualización periódica del año base del IPC y un cambio en la metodología del mismo¹, de tal manera que este esté en concordancia con las recomendaciones llevadas a cabo por la Organización Internacional de Trabajo (OIT)² y permita con ello la obtención de un indicador macroeconómico que garantice los fines planteados por esta institución.

Esta institución plantea que dentro de los usos que se puede dar al IPC se encuentra la indexación de valores monetarios³ tales como: salarios, pensiones, intereses, impuestos, rentas, etc. Por esta razón el correcto pronóstico de esta variable puede permitir una mejor planeación a futuro a sectores de la economía tan diversos como: la industria manufacturera, mercado de seguros y fianzas o los diferentes organismos estatales como ministerios o empresas públicas. Todos estos sectores requieren una previsión exacta de este indicador de manera tal que se les permita establecer de manera técnica las diferentes políticas de compensaciones, ajustes, seguros, asignación presupuestaria, inversiones, etc. Que adoptaran a futuro, optimizando de esta manera una importante cantidad de recursos.

El IPC también puede ser utilizado como una medida de comparación de dinamismo económico y paridad de poder adquisitivo entre dos o más

¹ Para mayor detalle de la metodología aplicada por el INEC, se recomienda revisar “La nueva metodología del índice de precios al consumidor (IPC) del Ecuador (base anual: 2014=100)”, cuyo identificador de dirección electrónica es: LC/W.731

² La OIT en colaboración con el Fondo Monetario Internacional (FMI), Organización de Cooperación y Desarrollo Económico (OCDE), Oficina Estadística de las Comunidades Europeas (Eurostat), Naciones Unidas y El Banco Mundial desarrollaron un manual del IPC denominado “Manual del Índice de Precios al Consumidor, teoría y práctica” cuyo ISBN es: 92-2-113699-X

³ Indexación se define como “un proceso en el cual el valor monetario de pagos o existencias se incrementan o reducen en proporción al cambio en el valor en determinado índice de precios” (ILO/IMF/OECD/UNECE/Eurostat/The World Bank, 2004, p.42)

economías (ILO/IMF/OECD/UNECE/Eurostat/The World Bank, 2004). Este aspecto resulta de vital importancia para una economía dolarizada como la Ecuatoriana, ya que puede ser utilizada por los inversores extranjeros como una forma de medir la salud del mercado en el que van a invertir, encontrándolo de esta manera más o menos atractivo para los negocios. Es decir que para el Ecuador, el poder prever esta variable le permite establecer políticas de ajuste de la misma en caso de que estas sean necesarias, de manera tal que beneficie a los objetivos de atracción de capitales y con ello brindar liquidez a la economía, evitando así el poner en riesgo a la dolarización.

El uso del IPC no se restringe solo a las autoridades económicas, por lo contrario, el conocimiento sobre este indicador está bastante generalizado en gran mayoría de la población, misma que mantienen un gran interés en permanecer actualizada sobre la evolución del nivel de precios de la economía y que por lo tanto forma parte activa de sus decisiones de consumo tanto presente como futuro (ILO/IMF/OECD/UNECE/Eurostat/The World Bank, 2004). Uno de los sectores económicos para el que resulta de gran relevancia el comportamiento de los precios en diferentes periodos de tiempo es el sector empresarial, el cual utiliza el IPC como un indicador para poder determinar los posibles costos por mano de obra durante cierto periodo, pronosticar posibles retornos por sus inversiones en diferentes periodos de tiempo, así como también le permite a las empresas obtener un bosquejo general del estado de la economía y en base a estas adaptar sus expectativas a futuro (Instituto Nacional de Estadísticas y Censos, 2015).

Por todos estos motivos la constante actualización y presentación de los resultados de este indicador por parte del Banco Central del Ecuador (BCE), así como el desarrollo de correctos modelos capaces de ofrecer pronósticos acertados a futuro, han resultado de mucho interés, tanto para individuos como para instituciones ya sea que estas se desenvuelvan en el sector público o privado. Dentro de los modelos utilizados tradicionalmente para el pronóstico de series de tiempo económicas destacan la familia de modelos ARIMA y VAR. Sin embargo el desarrollo que el campo de la informática ha venido experimentando desde la década de los 80 ha permitido que modelos cuyos

elevados requerimientos para el procesamiento de información resultaban restrictivos, empiecen a tener cada vez mayor inmersión en varias aplicaciones de las diferentes ciencias, como lo son los modelos de redes neuronales artificiales (RNA) (Chen, 2001).

Previo a la modelación multivariada del IPC es necesario determinar cuáles son los factores que más influyen en las fluctuaciones del mismo dentro de la economía. John Maynard Keynes en su libro *“The General Theory of Employment, Interest and Money”* aborda la inflación como un fenómeno producto del incremento de la demanda efectiva por sobre la oferta⁴ (Keynes, 1936). Por otro lado, dentro del enfoque clásico la inflación es abordada como un fenómeno que mantiene una relación directa con el flujo monetario⁵ (Fisher, 1922). La disputa ideológica entre estos dos enfoques daría origen al enfoque monetarista de la inflación, enfoque en el cual Milton Friedman afirma que la inflación siempre es consecuencia de fenómenos monetarios, sin embargo considera las aportaciones de Keynes en cuanto a que la demanda de dinero y los rendimientos de los activos financieros⁶ (Friedman, 1970).

Pese a todas las aproximaciones teóricas existentes respecto a la inflación, es menester de cada economía el estructurar aquellos modelos matemáticos, estadísticos o econométricos que mejor se adapten a sus particularidades, en especial cuando la economía estudiada presente singularidades que la

⁴ Actualmente se utiliza más frecuentemente la demanda agregada, dada por el consumo de los hogares y privados (C), inversión privada (I), el gasto público (P) y el saldo neto de exportaciones menos importaciones (X-M) de la siguiente manera: $DA = C + I + P + (X - M)$

Keynes indica que todo incremento sobre la demanda efectiva dentro de una economía que no se encuentre en pleno empleo, tendrá un efecto en el incremento de los precios.

⁵ Fisher realiza la formulación de la ecuación de cambio de la siguiente manera: $P * Q = M * V$

donde:
 P : nivel de precios
 Q : nivel de producción
 M : Cantidad de dinero
 V : velocidad de circulación de dinero

Dado que la formulación de Fisher parte desde un enfoque neoclásico, el nivel de producción es el de pleno empleo, es decir se asume constante. De la misma manera la velocidad de circulación de dinero es asumida constante en el corto plazo ya que esta es reflejo de los hábitos propios de los individuos que solamente cambian con el tiempo. Por esta razón el cambio en los precios esta dado solamente por cambios en las cantidades de dinero de una economía, mismo que puede ser aproximado a través de la cantidad de dinero en circulación que tienen los diferentes Bancos Centrales.

⁶ Milton Friedman reformula la teoría cuantitativa del dinero tomando en cuenta las críticas realizadas por Keynes como el efecto que la expansión monetaria tienen en el corto plazo.

desmarcan del tratamiento normal que se pueda dar a otras economías, como es el caso de la economía ecuatoriana, la cual al no contar con moneda propia⁷, tiene serias restricciones en cuanto a la aplicación de políticas de carácter monetario⁸.

En el año 2016, Cristina Carrera en su tesis doctoral "*The inflation – output GAP dynamics in Ecuador*" analiza la dinámica del crecimiento del PIB y de la inflación en el Ecuador, así como los factores determinantes de la misma para el periodo 1970-2014. En dicha tesis doctoral Carrera encuentra que variables como la inestabilidad política, episodios económicos previos marcados ya sea por una inflación galopante, desinflación o deflación, precios de los commodities⁹ y las políticas gubernamentales como expansiones monetarias agresivas, incremento o reducción del gasto público, creación o modificación de subsidios, ejercen presión sobre la inflación¹⁰ (Carrera, 2016).

Guerrero (2017) lleva a cabo un estudio donde busca determinar los factores que influyen en la inflación de una economía dolarizada así como la determinación de correctas políticas públicas enfocadas al control de la misma. En dicho estudio Guerrero utiliza un modelo de vectores auto regresivos (VAR) y un modelos de corrección de errores (VEC) para encontrar evidencia empírica sobre el efecto que variables macroeconómicas como el índice de precios de las materias primas, índice de tipo de cambio real (ITCR), índice de desarrollo de la actividad económica coyuntural (IDEAC), oferta monetaria (M1), liquides total (M2) y la deuda y gasto público como porcentaje del PIB tienen sobre la inflación. Guerrero encuentra que la variables más importantes para explicar el comportamiento de la inflación durante este periodo son la liquides total y los resultados fiscales (Guerrero, 2017).

1.1. Revisión de la literatura

⁷ La economía del Ecuador perdió su moneda propia (Sucre) desde el año 2000, año en el cual adoptó el dólar estadounidense como moneda de uso oficial.

⁸ La emisión monetaria es considerada una de las principales herramientas de política monetaria.

⁹ Commodities se definen como los precios de las materias primas o productos poco industrializados, los mismos que representan gran parte del porcentaje del total de producción exportada por el país.

¹⁰ Los mecanismos a través de los cuales estas variables afectan la inflación son estudiados más profundamente en el trabajo de Cristina Carrera titulado: "*The inflation – output GAP dynamics in Ecuador: A Descriptive Analysis during Episodes of Positive and Negative Output Gaps. 1970 – 2014*".

Al igual que en todas las ciencias, dentro del campo de la economía se ha buscado desarrollar modelos de series de tiempo cuyos resultados sean eficientes y confiables. Dichos modelos son aplicados mayoritariamente en la previsión de series económicas (especialmente financieras) y en el análisis de políticas económicas, sea este el efecto directo que tienen una variable explicativa sobre la variable de interés o sea el predecir el efecto que el shock en una variable explicativa podría tener sobre la variable de interés (Echevarría, 1975) ¹¹.

Con esta finalidad, los especialistas en modelación económica han recurrido en mayor medida a los modelos de series de tiempo auto regresivos como lo son los modelos ARIMA y VAR¹², mismos que han ofrecido resultados eficientes en el pronóstico de las diferentes series económicas. Sin embargo el carácter interdisciplinario de las actuales instituciones inmersas en la investigación, junto con el desarrollo de las tecnologías informáticas para el procesamiento de información, han permitido que herramientas antes limitadas a su uso solo por parte expertos en ciertas áreas, se expandan hacia las demás ciencias con resultados interesantes. Una de estas herramientas son las llamadas redes neuronales artificiales, mismas que durante los últimos 10 años han visto un resurgimiento en su aplicabilidad en varias disciplinas, en gran parte motivados por el creciente interés en el desarrollo de la inteligencia artificial (Tostado, Orneras, Espinel, Puga, 2015).

Iqbal y Naveed (2016) utilizan la aproximación de Box-Jenkins¹³ para llevar a cabo la comparación del desempeño que diferentes modelos ARIMA tienen en el pronóstico del IPC de Pakistán en el periodo 1970-2016. Iqbal y Naveed

¹¹ Echevarría Indica que en la creación de modelos de pronósticos se diferencian claramente en 2 tendencias, por un lado quienes creen que los modelo univariantes son modelos excelentes para fines de pronóstico ya que se captura la evolución de la serie, por el otro aquellos que creen que las fluctuaciones en una variable corresponden a shocks exógenos a la misma, por lo que estos cambian en el tiempo y se requiere de modelos multivariantes capaces de capturar estos shocks y ofrecer mejores pronósticos (Echevarría, 1975).

¹² En realidad se trata de una familia de modelos ARIMA y VAR, cada uno con diferencias que les permite ajustarse de mejor manera a las series destinadas a modelar.

¹³ Box-Jenkins proponen una metodología para la selección del modelo ARMA o ARIMA que mejor resultado de pronóstico ofrezca, metodología que consta de 3 pasos: Identificación y selección del modelo, estimación de los parámetros y la comprobación del ajuste del modelo seleccionado (Box, Jenkins, Reinsel 1994).

encuentra que la serie es integrada de primer orden, en cuanto al esquema ARMA que mejores resultados les brindo en el pronóstico del IPC fue el esquema ARMA(10,1)¹⁴.

Dentro del entorno nacional, Erraez (2005) realiza un análisis del proceso inflacionario en la economía ecuatoriana en los primeros años post dolarización, para dicho análisis Erraez se vale de los modelos VAR y ARIMA con la finalidad de determinar las variables que pueden afectar la inflación nacional y realizar la proyección de la inflación para el año 2005, adicionalmente se buscan posibles quiebres estructurales durante los primeros tres años de dolarización y el periodo posterior al año 2004 (Erraez, 2005). Los resultados obtenidos por Erraez indican que el modelo que mejor desempeño tiene para el pronóstico es un esquema ARMA (2,2) integrado de primer orden¹⁵.

Otra familia de modelos ampliamente utilizada dentro de la disciplina económica son los modelos VAR, los cuales han mostrado ser bastante eficiente para realizar pronósticos a mediano plazo (Sims C. , 1986)¹⁶. A diferencia de los modelos de ecuaciones simultaneas, el gran atractivo de los modelos VAR es que trata cada variable que conforma el sistema de manera simétrica, por lo tanto se evita el problema de asegurar la exogeneidad de la variable de interés. Sims (1980) destaca el grave error que representa el hecho de tomar variables solamente como endógenas y no como exógenas para mantener una correcta especificación del modelo dentro de las series económicas, mismas que por su naturaleza tienden a relacionarse la una con la otra en ambas direcciones.

Para el caso de la economía ecuatoriana, Morán (2014) realiza un análisis econométrico sobre los determinantes de la inflación en el Ecuador, para dicho

¹⁴ Iqbal y Naveed realizan las pruebas de diagnóstico en 2 modelos tentativos: $Y_t = \alpha_0 + \alpha_1 X_{t-4} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$ y $Y_t = \alpha_0 + \alpha_1 X_{t-4} + \alpha_2 X_{t-10} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$ (Iqbal M. y Naveed A., 2016).

¹⁵ A partir de este punto se utilizara la nomenclatura técnica para los modelos ARIMA (o ARMA en caso de que el orden de integración sea 0).

¹⁶ Dentro de su paper "Are forecasting models usable for policy analysis" Sims lleva a cabo un análisis de las bondades que los modelos VAR (hasta ese entonces usado ampliamente para realizar pronostico) tienen en temas de política económica.

análisis Morán utiliza un modelo VAR para determinar la relación entre variables correspondientes a IPC, salarios, M1, tipo de cambio real efectivo y el gasto público. Dentro de este estudio se encontró que el modelo VAR es $I(1)^{17}$, con 10 rezagos seleccionados en base al comportamiento de los residuos, pese al número de rezagos utilizados por Morán, el modelo seleccionado presenta correlación serial y no normalidad.

El número de rezagos que tiene el modelo VAR propuesto por Morán, junto con la cantidad de variables con las que se trabaja, conlleva a que el número de parámetros a estimar haya sido muy elevado y para el caso de la inflación ecuatoriana, economía de la cual se dispone solamente de información a partir del año 2000¹⁸ representa un serio problema. Este elevado número de parámetros a estimar puede llevar a incrementar el error en las estimaciones y con ello obtener pronósticos menos acertados. Con la finalidad de reducir este error en el pronóstico es recomendable realizar pruebas de causalidad previas al modelo VAR, de manera que se permita trabajar con un número reducido de variable, pero que dichas variables resulten ser las más significativas (Guzman y Alba, 2008).

Tal como lo señala Nakamura (2004) dentro de la ciencia económica los modelos de RNA han tenido un mayor uso para el pronóstico de tipo de cambio y de series financieras, mientras que el modelado multivariante de series macroeconómicas a través de RNA no ha sido objeto de estudios más profundos, en parte debido al limitado uso que los investigadores hacen de técnicas de pre procesado y detención temprana¹⁹ al utilizar una RNA, razón por la cual los resultados obtenidos no siempre son eficientes (Nakamura, 2004). De la totalidad de modelos que han sido desarrollados dentro de la familia de las redes neuronales artificiales, las topologías de red de tipo perceptrón multicapa han sido los más utilizados por las ciencias económicas, esto debido a la calidad obtenida en los pronósticos de las diferentes series a

¹⁷ Integrado de primer orden.

¹⁸ El utilizar los datos de la inflación correspondiente al periodo previo a la dolarización tergiversaría los resultados obtenidos, debido a la evidente diferencia entre la realidad económica del Ecuador en el periodo pre dolarización y post dolarización.

¹⁹ Mismas que serán tratadas con mayor detenimiento en las siguientes secciones del estudio.

través del uso de una arquitectura de red relativamente simple (Velazquez y Franco, 2010).

Las arquitecturas de red de tipo perceptrón multicapa, también conocidas como redes “*back propagation*” han demostrado ciertas fortalezas en el pronóstico de series de tiempo tales como: la capacidad de modelar no linealidades, alta capacidad de adaptabilidad y aprendizaje, la no necesidad de una forma funcional específica, entre otras (Wang, Zeng y Chen, 2014). A su vez Wang et al. (2014) realizan un análisis del problema de mínimos locales y proponen algoritmos alternativos al descenso del gradiente²⁰, además los autores prueban la validez de la arquitectura con una sola capa oculta como aquella arquitectura que permite aproximar de mejor manera cualquier función no lineal.

Esta capacidad para modelar no linealidades junto a la gran adaptabilidad que tienen los modelos de RNA a shocks que puedan ocasionar cambios estructurales en la variable de estudio (ya sean cambios en media o varianza), los convierten en modelos muy atractivos para realizar pronósticos a horizontes de 12 meses²¹ o superiores, ya que tienen la capacidad de captar y generalizar cualquier anomalía en la serie de manera muy rápida (Laptev, Yosinski, Erran, Smyl, 2017).

En cuanto a la literatura existente sobre el desarrollo de modelo de RNA para el pronóstico de la inflación²², McNelis y McAdam (2004) aplican modelos lineales y de RNA para llevar a cabo el pronóstico de la inflación de 3 economías²³ y realizan una comparación de dichos resultados. Para llevar a cabo dicha comparación, McNelis y McAdam utilizan modelo lineales de media recortada, mientras que para el modelo RNA siguen un algoritmo back propagation con el uso de una función de activación lineal, para la estimación de los pesos neuronales y los límites de la función de activación de cada neurona los autores recurren al procedimiento de aproximación híbrida de Quagliarella y

²⁰ Wang et al. (2014) proponen como algoritmos de entrenamiento alternativos como: métodos algebraicos lineales para la selección de pesos iniciales, algoritmos de evolución inteligente tales como algoritmo genético o la optimización de enjambre de partículas.

²¹ Por lo general el horizonte de interés para los hacedores de política es de 12 meses.

²² Varios autores toman a la variación del IPC como una muy buena aproximación de la inflación.

²³ Estados Unidos, Japón y la zona euro.

Vicini²⁴, con la finalidad de evitar mínimos locales (McNelis y McAdam, 2004). Los autores obtienen evidencia a favor del uso de modelos de RNA, sobre todo debido al mejor desempeño que estos modelos muestran en periodos de incertidumbre o cambios estructurales.

Para la evaluación de la calidad de pronóstico de cada modelo, dado que el objetivo de estos es minimizar el error de las predicciones, el criterio más importante en la selección y evaluación de estos modelos es el error cuadrático medio (Pindyck y Rubinfeld, 2001), no obstante existen una serie de indicadores de la calidad de pronóstico que pueden brindar información adicional sobre el desempeño de dichos modelos²⁵.

En el caso de que el modelo de pronóstico tenga fines de apoyo a la política económica, a más de cumplir con todos los supuestos establecidos para la validación de los mismos, se debe asegurar que el pronóstico del modelo sea estable en el tiempo, es decir que este modelo mantenga un buen desempeño ante el ingreso de nueva información. Con este fin Salazar (2009) realiza una evaluación del desempeño de los modelos de redes neuronales ante el ingreso de nueva información, para ello Salazar se basa en la metodología de estimación recursiva de parámetros por fuera de muestra (conocida como “*rolling*”), Salazar concluye que la existencia de no linealidades dentro de las series tratadas hace que los modelos de RNA superen ampliamente al modelo ARIMA, para llegar a esta conclusión se utiliza la técnica de rolling con 8 iteraciones a un horizonte de pronóstico de 12 meses y se compara la raíz del error cuadrático medio en cada iteración.

²⁴ La aproximación híbrida de Quagliarella y Vicini propone que para obtener el conjunto de pesos $[\omega \ \gamma \ \beta]$ de la red, previo a la fase de entrenamiento de la misma, se deben estimar a través de un algoritmo iterativo un conjunto de pesos iniciales $[\hat{\omega}_0 \ \hat{\gamma}_0 \ \hat{\beta}_0]$, este vector de pesos se utilizan como valores iniciales para la minimización del descenso del gradiente, algoritmo que nos dará el vector de pesos final de la red para realizar los pronósticos (McNelis y McAdam, 2004).

²⁵ Indicadores tales como la raíz del error cuadrático medio (RECM), error porcentual promedio (MPE), error absoluto porcentual promedio (MAPE), U de Theil.

1.2. Metodología

En el análisis de modelos de series de tiempo con fines de pronóstico existen amplios estudios para el correcto tratamiento de las series²⁶, en todos ellos se destaca la importancia que el análisis previo de la serie tiene al momento de elegir el modelo de pronóstico. El análisis de la serie nos permite obtener una noción del comportamiento que ha tenido la misma en el tiempo y con ello identificar la presencia ya sea de uno o varios componentes de tendencia, componentes cíclicos, componentes estacionales o componentes aleatorios característicos de las series temporales.

Un concepto importante para la modelación de series de tiempo de carácter macroeconómico es el de “proceso estocástico estacionario”, en este aspecto Gujarati y Porter (2009) señalan que *“se dice que un proceso estocástico es estacionario si su media y su varianza son constantes en el tiempo y si el valor de la covarianza entre dos periodos depende sólo de la distancia o rezago entre estos dos periodos, y no del tiempo en el cual se calculó la covarianza”* (p.740)²⁷. Este concepto adquiere especial relevancia para aquellos modelos destinados para pronóstico, ya que se requiere estacionariedad en la serie para poder asegurar que los resultados obtenidos en la estimación pueden ser generalizados para otros periodos de tiempo y no se trate de relaciones espuria.

²⁶ Enders (2009) analiza de manera profunda todos los aspectos que se deben tener en cuenta en el análisis de una serie de tiempo previo a la selección del modelo que se decida utilizar, así como la corrección de posibles problemas que se presenten. Gujarati y Porter (2009) tratan ampliamente los temas de multicolinealidad, heterocedasticidad y auto correlación en los cap. 10, 11 y 12 respectivamente de su libro *“Econometría 5ta edición”*, así como una muy buena aproximación a modelos con fines de pronóstico. Finalmente Stock y Watson (2012) desarrollan de manera detallada el uso de modelos con fines de pronóstico. Dichos libros sirvieron de fundamento en el tratamiento de las series y la selección de los modelos que se desarrollan en este documento, por lo que se puede referir a ellos si se desea profundizar en los conceptos sobre series de tiempo tratados en esta investigación.

²⁷ La estacionariedad a la que Gujarati y Porter hacen referencia es a la estacionariedad débil, siendo esta requisito suficiente para los modelos de pronóstico. Se define que una serie es débilmente estacionaria cuando: $E(Y_t) = \mu$; $var(Y_t) = E(Y_t - \mu)^2 = \sigma^2$; $cov(Y_t) = E[(Y_t - \mu)(Y_{t+k} - \mu)] = \gamma^2$

Es muy común que las series de carácter macroeconómico resulten ser un proceso de tipo caminata aleatoria²⁸, donde la variable en el periodo t toma el valor de la variable en el periodo $t-1$ más un shock aleatorio:

$$Y_t = Y_{t-1} + u_t \quad \text{con} \quad u \sim N(0, \sigma^2)$$

La ventaja de los procesos de este tipo es que se consigue la estacionariedad en los mismos a través de un proceso conocido como diferenciación, que no es más que a Y_t restarle Y_{t-1} . En la mayoría de series económicas se ha visto que solamente se necesita de una diferencia para conseguir que la variable sea estacionaria²⁹ y dado que este es requisito fundamental para la eficiencia de los modelos de pronóstico, todas las variables fueron trabajadas en primera diferencia³⁰ tanto para los modelos paramétricos como para los no paramétricos.

Para la detección de la estacionariedad los métodos gráficos nos pueden dar una idea de si la serie es o no estacionaria, otra prueba grafica de estacionariedad es el análisis de la función de autocorrelación (FAC). En cuanto a las pruebas formales, la más utilizada es la prueba de raíz unitaria (también conocido como test Dickey-Fuller) y test Dickey – Fuller aumentado³¹.

1.2.1. Modelo autoregresivo integrado de media móvil (ARIMA).

Por lo general los esquemas ARIMA (p, d, q) utilizados en series económicas resultan ser integradas de orden uno [$I(1)$], razón por la cual la preocupación fundamental recae sobre la selección del esquema ARMA más adecuado, un esquema ARMA se encuentra definido de la siguiente forma:

²⁸ Un proceso de caminata aleatoria para $t = 1$ se define como: $Y_1 = Y_0 + u_1$, de la misma manera para $t=2$ se tiene que $Y_2 = Y_1 + u_1$. Sin embargo $Y_1 = Y_0 + u_1$, entonces $Y_2 = Y_0 + u_1 + u_2$. Se puede deducir entonces que $Y_t = Y_0 + u_1 + u_2 + \dots + u_t$, es decir que dentro de un proceso de caminata aleatoria el valor en el tiempo t es igual al valor inicial ($t=0$) más la suma de shocks desde Y_0 hasta Y_t , lo que implica que el proceso es no estacionario en varianza.

²⁹ La mayoría de series económicas muestran ser integradas de orden uno [$Y_t \sim I(1)$].

³⁰ En la sección Anexos B se aprecia las variables tanto en niveles como en primera diferencia.

³¹ Enders (2009), Gujarati y Porter, Stock y Watson desarrollan de manera detallada cada uno de los test de estacionariedad utilizados tradicionalmente.

$$Y_t = \theta + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \beta_0 u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \dots + \beta_q u_{t-q}$$

Donde:

- $\alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p}$ corresponde al proceso AR(p) de tipo $Y_t = \alpha_0 + \sum_{i=1}^p \alpha_i Y_{t-i} + u_t$.
- $\beta_0 u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \dots + \beta_q u_{t-q}$ corresponden a un proceso MA(q) de tipo $Y_t = \mu + \sum_{i=0}^q \beta_i u_{t-i}$.

Para la selección del esquema ARIMA (p, d, q) que mejores resultados de pronóstico ofrezca, generalmente se utiliza la metodología propuesta por Box – Jenkins (BJ) misma que consta de 4 pasos, tal como se puede apreciar en el esquema representado en el grafico 1.

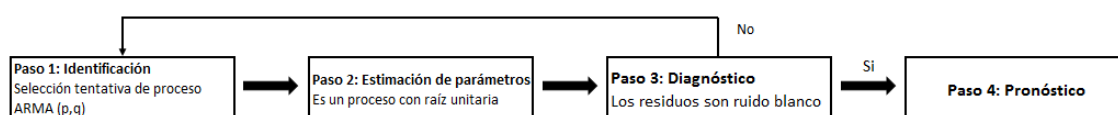


Gráfico 1: Metodología Box – Jenkins para la selección del esquema ARIMA (p, d, q)

Fuente: Gujarati y Porter (2010)

Elaboración propia

Es común que varias series económicas presenten componentes estacionales, series como la producción agrícola, niveles de ventas e inclusive el IPC puede presentar estos componentes producidos en los meses del año que más consumo existe, como las fechas de navidad, día de las madres, días feriados, etc. El tratamiento ante la presencia de componentes estacionales no difiere del tratamiento normal del esquema ARMA (p,q), siendo necesario la introducción de un componente que capture la estacionalidad, la cual se puede detectar fácilmente a través del análisis del grafico de AC y PAC³². De esta manera, para un esquema ARIMA (1,1,1) que presenta un componente estacional $t = 12$, el modelo se convierte entonces en un esquema SARIMA con la siguiente forma: $Y_t = \theta + \alpha_1 Y_{t-1} + \alpha_{12} Y_{t-12} + \beta_1 u_{t-1}$.

³² El IPC en tratamiento cuenta con un componente claramente estacional, a razón de ejemplo el AC y PAC de la serie puede apreciarse en la sección Anexos C, subsección Anexos C.1, Grafico Anexo 16.

1.2.2. Modelo de vectores autoregresivos (VAR)

Como se mencionó anteriormente, los modelos VAR tratan a cada variable que conforma el sistema de manera simétrica, así para un sistema VAR conformado por dos variables³³, su representación está dado por:

$$Y_t = \beta_{10} + \sum_{j=1}^p \beta_{1j} Y_{t-j} + \sum_{j=1}^p \gamma_{1j} Z_{t-j} + u_{1t}$$
$$Z_t = \beta_{20} + \sum_{j=1}^p \beta_{2j} Y_{t-j} + \sum_{j=1}^p \gamma_{2j} Z_{t-j} + u_{2t}$$

Como se puede apreciar claramente en el sistema VAR representado anteriormente, el modelo tiene tantas ecuaciones como variables conformen el mismo, también se puede apreciar como dentro del sistema las variables explicativas en cada ecuación son los respectivos rezagos de las demás variables y de sí misma, por lo que la cantidad de coeficientes a estimar se multiplica con la introducción de cada variable y sus respectivos rezagos³⁴. Este alto número de coeficientes adicionales a estimar con la inserción de cada nueva variable, provoca que sistemas VAR con muchas variables vean un significativo incremento en el error de predicción³⁵.

Con la finalidad de asegurar un correcto desempeño del modelo VAR, así como la fiabilidad que ofrezca el mismo como modelo de soporte para política económica, todas las variables que forman parte del mismo deben ser integradas del mismo orden, los residuos deben ser ruido blanco y u_{1t} y u_{2t} deben estar incorrelacionados³⁶. En cuanto a la selección del número ideal de rezagos del modelo, generalmente se utilizan los criterios de información de

³³ El procedimiento para un sistema VAR con k variables resulta análogo al esquema VAR de dos variables

³⁴ Stock y Watson (2012) ejemplifican que para un modelo VAR con 5 variables y 4 rezagos tendrá 21 coeficientes por cada ecuación, para un total de 105 coeficientes a estimar en el sistema total.

³⁵ No está demás mencionar que las implicaciones prácticas para investigadores que trabajen con modelos VAR es el especial cuidado que deben tener en la selección y tratamiento de las variables del sistema.

³⁶ Enders (2009) trata ampliamente las implicancias de estos supuestos y la importancia de la estacionariedad en los modelos VAR.

Akaike o de Schwarz, con la finalidad de evitar problemas de multicolinealidad o mala especificación por la utilización de un número errado de rezagos.

1.2.3. Modelo de redes neuronales artificiales (RNA)

La inteligencia artificial (IA) ha tenido un gran avance durante los últimos 10 años, en gran parte impulsada por la fuerte inversión que importantes empresas tecnológicas han realizado en esta área. Actualmente diferentes componentes con cierto grado de inteligencia artificial se encuentran en áreas tan diversas como las aplicaciones móviles, los asistentes virtuales, videojuegos, modelos dinámicos no paramétricos para la detección de patrones, aplicaciones de bolsas de valores, etc. Sin embargo los campos de investigación sobre los que se trabaja resulta ser mucho más amplio y con importantes implicancias económicas tales como: desarrollo de vehículos autónomos, desarrollo de inteligencia artificial completamente autónoma, asistentes virtuales autónomos, tecnificación de industria manufacturera, entre otras. Es por esto que el campo de la IA tiene un futuro promisorio y se apuntala como uno de los factores más importantes para el desarrollo de las industrias del futuro y con ello de las economías que hayan invertido en la investigación de estas tecnologías³⁷.

Parte importante del desarrollo de la inteligencia artificial es el conocer cómo funciona el cerebro humano y cuáles son los componentes y procesos que le permiten tener un complejo sistema de razonamiento, sobre todo los procesos de abstracción y generalización. Las RNA son el producto de la investigación de estos procesos y el intento de la recreación de los mismos a través de algoritmos desarrollados en lenguaje matemático.

La capacidad de razonamiento del ser humano nace de un complejo sistema de interconexiones entre células llamadas neuronas, son los procesos electroquímicos producidos entre estas neuronas y la cantidad de conexiones

³⁷ Russell y Norvig (2010) así como Goodfellow, Bengio y Courville (2016) dedican el primer capítulo de su libro a destacar los avances que muestran estas aplicaciones informáticas y el resurgimiento del uso de esta herramienta, de igual manera existen portales web de ciencia y tecnología que mantienen una actualización constante sobre este campo. A manera de referencia se sugiere la agencia española para la inteligencia artificial (AEIPA) cuya dirección electrónica es <http://www.aepia.org/aepia/>.

existentes entre las mismas las que definen la capacidad de razonamiento que tiene una persona. Una simplificación matemática de estos procesos se observa en el grafico 2.

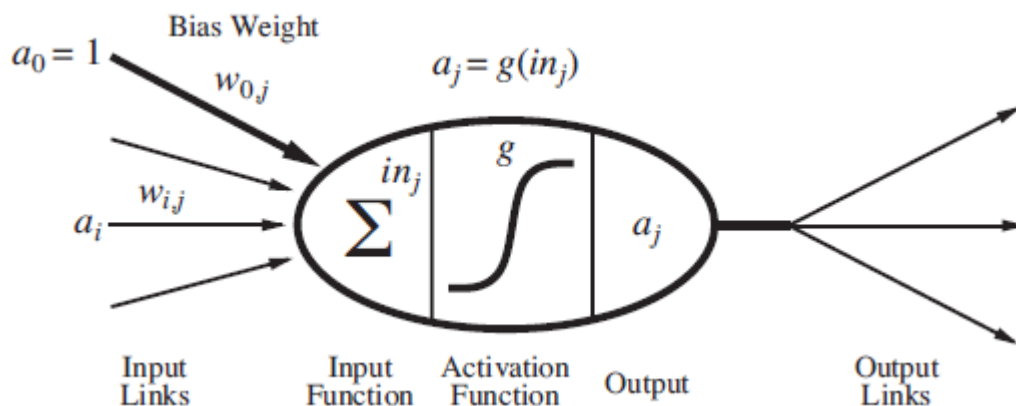


Gráfico 2: Simplificación matemática de una neurona (neurona artificial)

Nota. Recuperado de "Artificial intelligence, a modern approach, 3rd edition", Russell, S., Norvig, P., 2010, p. 728, New Jersey, Estados Unidos: Pearson education Inc.

Donde:

a_i y w_{ij} son los valores de entrada (variables explicativas) y pesos sinápticos respectivamente³⁸. En el grafico se ejemplifican las interconexiones existentes en una neurona ya que como se puedes observar, la flecha $a_i w_{ij}$ corresponde al nodo de los valores de entrada, mientras que las demás flechas representan las demás señales que recibe esta neurona, que son solamente las señales de salida de las otras neuronas que componen la RNA.

Σin_j es la regla de propagación, que para el caso del problema de estudio, resulta ser la combinación lineal de los valores de entrada a_i y w_{ij} de la siguiente forma:

$$h_i(a_1, a_2, \dots, a_n; w_{i1}, w_{i2}, \dots, w_{in}) = \sum_{i=1}^n w_{ij} a_i$$

³⁸ Los pesos sinápticos son valores numéricos constantes análogos al grado de sinapsis de las neuronas, estos valores numéricos tienen la capacidad de incrementar o reducir la "importancia" que tiene un valor de entrada dentro de la regla de propagación, así como definir si su regla de propagación es inhibidora (peso negativo) o excitadora (peso positivo).

a_0 y w_{0j} son valores de activación o umbrales³⁹ que por lo general se resta del total pos sináptico de la regla de propagación o función de entrada de la siguiente manera:

$$h_i(a_1, a_2, \dots, a_n; w_{i1}, w_{i2}, \dots, w_{in}) = \sum_{i=1}^n w_{ij}a_i - w_{0j}a_0$$

g_j es la función de activación asociada a la neurona, esta es la parte del algoritmo encargada de transformar la función de entrada en la respuesta de la neurona a_j como se muestra en la siguiente expresión:

$$a_j = g(h_i) = g_i \left(\sum_{i=1}^n w_{ij}a_j - w_{0j}a_0 \right)$$

A la función de activación se la conoce también como función limitadora debido a que precisamente “limita” la función de entrada a un rango permisible de valores, los cuales son definidos por la función de activación que se utilice.

Las funciones de activación que se utilizan con más frecuencia en las RNA con fines de pronóstico son:

Función de activación lineal: función de la forma $f(x) = x$, esta función se caracteriza por no tener límites para el rango de salida a_j , aunque resulta conveniente para la aplicación de esta función la normalización previa de las series trabajadas.

Función de activación sigmoidea logarítmica: función de la forma $f(x) = \frac{1}{1+e^{-x}}$, es decir que esta función da un rango de salida a_j comprendido entre 0 y 1. Lo importante de esta función es que resulta diferenciable a lo largo de todo su dominio, mismo que es requisito para la mayoría de los algoritmos de aprendizaje utilizados para fines prácticos que serán tratados más adelante.

Función de activación tangente hiperbólica: función de la forma $f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$, esto implica que el rango de valores permisibles para a_j se

³⁹ En analogía a los modelos econométricos, estos valores corresponden a los términos de intercepto (β_0)

encuentra comprendido entre -1 y 1, al igual que la función sigmoidea logarítmica, esta también resulta ser diferenciable a lo largo de todo su dominio.

Con la finalidad de optimizar los resultados obtenidos por la RNA es conveniente realizar una transformación de las series trabajadas, de manera tal que se encuentren dentro del rango de valores permisibles de cada función de activación.

Una de las transformaciones más sencillas y que ofrece buenos resultados son la estandarización estadística dada por:

$$a_s = \frac{a_i - \bar{a}_i}{\sigma_{a_i}}$$

Para el caso de las redes donde se opte por la función sigmoidea logarítmica o tangente hiperbólica se escala las series para que queden delimitadas entre 0 y 1 o -1 y 1 según sea el caso, el escalonamiento de la serie se realiza de la siguiente manera:

$$a' = \frac{(a_i - a_{i \min})(ls - li)}{a_{i \max} - a_{i \min}} + li$$

El funcionamiento de las neuronas artificiales es muy similar al de las neuronas biológicas, de esta manera se tiene las variables explicativas a_i que alimentan la red, las cuales son ponderadas por un peso w_{ij} en función de si estas variables resultan ser excitadoras (w_{ij} positivo), inhibidoras (w_{ij} negativo) o no afectan a las demás capas (w_{ij} de cero). Estas variables ponderadas, así como los respectivos valores de activación de cada neurona forman la regla de propagación (valor numérico) la cual tras pasar por la función de activación arroja un resultado correspondiente a esa neurona, este resultado puede retroalimentar a las demás neuronas de la misma capa o a sí misma, servir de valor de entrada para las neuronas existentes en capas posteriores o ser un valor final de salida de la red, dependiendo de la topología de red adoptada.

Una vez definida la unidad de procesamiento de la red (neurona), el siguiente paso es definir la manera en la que estas neuronas artificiales interactúan

dentro de la misma, la topología de red más utilizada en el modelado de series económicas son las conocidas como perceptrón multicapa o debido a su algoritmo de aprendizaje también conocidas como redes back-propagation.

La fortaleza de esta topología de red es la inclusión de capas ocultas (como se aprecia en el gráfico 3) entre la capa de entrada y de salida, mismas que le permiten a la red crear zonas de discriminación en el hiperespacio a través de la generación de diversos hiperplanos. Estas redes generan zonas de exclusión mucho más precisas, lo que en fines prácticos se traduce en pronósticos con mejor ajuste.

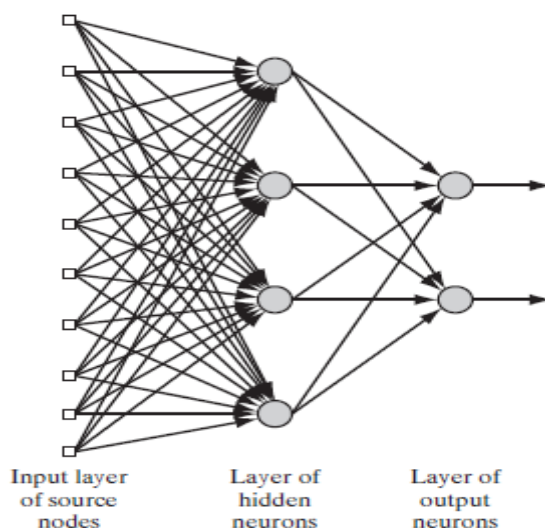


Gráfico 3: Topología perceptrón multicapa

Nota. Recuperado de “*Neural networks and learning machines, 3rd edition*”, Haykin, S., 2009, p. 22, Ontario, Canadá: Pearson education Inc.

La topología perceptrón multicapa resulta ser una red de conexión completa ya que como se ve en la gráfica 3 cada nodo de salida de las neuronas correspondientes a una capa i se conectan con los nodos de entrada de una capa j . Es importante destacar que esta topología cuenta con una o más capas ocultas, cada una compuesta por n neuronas ocultas, cada una con una función de activación diferenciable y pesos sinápticos w_{ij} propios para cada entrada a_i en cada capa j .

Para la fase de entrenamiento de la red⁴⁰ de tipo perceptrón multicapa, el algoritmo más utilizado es el de back-propagation, mismo que puede resumirse en dos etapas: una primera etapa en la cual se establecen los valores iniciales de los pesos sinápticos w_{ij} y se propagan las señales de salida hacia capas posteriores, es decir que en esta fase solamente se trabaja en los nodos de salida y en los valores umbral. En la segunda etapa se compara la salida de la red con los valores deseados, obteniendo así una señal de error la cual se propaga a través de la red en dirección contraria, en esta fase los pesos sinápticos son modificados de tal manera que se reduzca el error entre la salida de la red y el valor deseado.

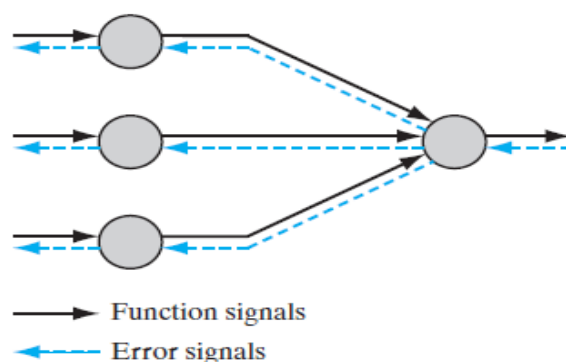


Gráfico 4: Funcionamiento del algoritmo back-propagation

Nota. Recuperado de “*Neural networks and learning machines, 3rd edition*”, Haykin, S., 2009, p. 125, Ontario, Canadá: Pearson education Inc.

El algoritmo back-propagation es un algoritmo de entrenamiento supervisado el cual consiste en minimizar una función de error $\varepsilon(w_{ij})$ con respecto a w_{ij} de manera tal que se encuentre una solución optima w_{ij}^* tal que satisfaga la condición $\varepsilon(w_{ij}^*) \leq \varepsilon(w_{ij})$.

Siendo $\varepsilon(w_{ij})$ una función continuamente diferenciable de un vector de parámetros w_{ij} , nos enfrentamos a un problema de optimización no restringido donde la condición de optimización necesaria está dada por:

$$\nabla \varepsilon(w_{ij}^*) = 0$$

Donde ∇ es el operador gradiente, por lo que $\nabla \varepsilon(w_{ij}^*)$ está definido por:

⁴⁰ Fase clave en el diseño de la red ya que se establecerán los valores umbrales y los pesos sinápticos que se utilizarán para realizar la generalización de la red.

$$\nabla \varepsilon(\mathbf{w}_{ij}^*) = \left[\frac{\partial \varepsilon(\mathbf{w}_{ij}^*)}{\partial w_1}, \frac{\partial \varepsilon(\mathbf{w}_{ij}^*)}{\partial w_2}, \dots, \frac{\partial \varepsilon(\mathbf{w}_{ij}^*)}{\partial w_n} \right]_j$$

Para llevar a cabo el proceso de optimización se recurre al método conocido como descenso pronunciado⁴¹, en este algoritmo se recurre al ajuste sucesivo del vector de pesos \mathbf{w} en el sentido apuesto al gradiente $\nabla \varepsilon(\mathbf{w}_{ij}^*)$ de la siguiente manera⁴²:

$$\mathbf{w}(t+1) = \mathbf{w}(t) - \eta \frac{\partial \varepsilon(\mathbf{w}_{ij}^*)}{\partial w_{ij}} = \mathbf{w}(t) - \eta \nabla \varepsilon(\mathbf{w}_{ij}^*)$$

Donde η es una constante positiva conocida como el ratio de aprendizaje y $\nabla \varepsilon(\mathbf{w}_{ij}^*)$ es el vector gradiente evaluado en el punto \mathbf{w}_{ij}^* . El algoritmo aplica la corrección η en cada iteración t de manera tal que se aproxime a la solución óptima, de este punto destaca la correcta selección de un ratio de aprendizaje η de tal manera que se evite el caer en mínimos locales por la selección de un η muy grande o una muy lenta convergencia hacia la solución óptima del algoritmo en caso de elegir un valor η muy pequeño⁴³. Este algoritmo se repite mientras no se obtenga un error deseado o se haya llevado a cabo el suficiente número de iteraciones, un método común utilizado para detener el algoritmo es la detención temprana que nos permite evitar mínimos locales que pueden traer consigo una mala generalización de la red.

Una variación a esta topología es la red neuronal recurrente en la cual el nodo de salida de la neurona a más de alimentar a las capas posteriores, también retroalimenta a la capa de la que forma parte, como se muestra en el grafico 4. La ventaja de esta topología de red es que le permite adaptarse de mejor manera a cambios bruscos en los datos, así como una mayor estabilidad ante posibles cambios estructurales que otras topologías de red utilizadas frecuentemente.

⁴¹ Este es el método más utilizado para la minimización del gradiente en aplicaciones económicas, sin embargo en Haykin (2009) se encuentran algoritmos alternativos como el método de Newton o la metodología Gauss-Newton, que se pueden utilizar en el algoritmo back-propagation.

⁴² Este es el planteamiento formal de la regla de corrección de error.

⁴³ Para una explicación matemática de los algoritmos utilizados, se puede referir al capítulo 3 y 4 de Haykin (2009).

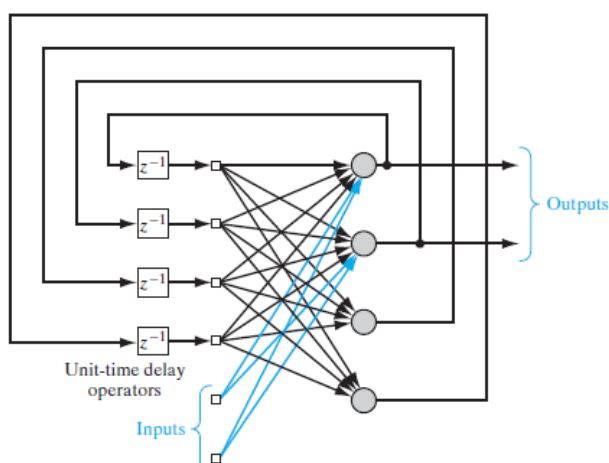


Gráfico 5: Red neuronal artificial recurrente.

Nota. Recuperado de “*Neural networks and learning machines, 3rd edition*”, Haykin, S., 2009, p. 24, Ontario, Canadá: Pearson education Inc.

1.2.4. Pronóstico

La calidad de cada modelo se evalúa a través de los pronósticos por fuera de muestra que estos realizan, para este fin se reservan los últimos 18 y 12 meses del IPC (dependiendo si se evalúa el pronóstico para el corto o mediano plazo), los diferentes modelos se estiman prescindiendo de estos datos, posteriormente se realiza un pronóstico del IPC para el periodo h y se contrasta dicho pronóstico con los valores reservados para la evaluación de los modelos⁴⁴. Debido a las transformaciones especificadas más adelante y la naturaleza propia de las series, se trabaja con 5 medidas estadísticas las cuales son: error cuadrático medio (ECM), raíz del error cuadrático medio, error porcentual promedio (MPE), promedio de error porcentual absoluto (MAPE), desviación del APE y U de Theil⁴⁵.

Para realizar los pronósticos, tanto para el modelo ARIMA como para el modelo VAR⁴⁶ se realizan predicciones multiperiodo iteradas (Stock y Watson, 2012) que no es más que estimar el modelo utilizando los datos hasta el periodo t y con este modelo se pronostica $t + 1$, para el pronóstico en $t + 2$ se reestima el modelo tomando en consideración el valor pronosticado de $t + 1$, el proceso es

⁴⁴ Los diferentes mecanismos de evaluación se pueden apreciar en la sección anexos A.

⁴⁵ Las particularidades de cada estadístico, condiciones para su aplicación, así como la interpretación de los mismos se puede encontrar de manera detallada en Vélez y Nieto (2016).

⁴⁶ A diferencia del modelo ARIMA, los modelos VAR realiza el pronóstico de todas las variables que componen el sistema.

el mismo hasta que se obtiene el horizonte de pronóstico h deseado. Para el pronóstico de los modelos RNA, una vez finalizado el proceso de entrenamiento y establecidos los pesos w_{ij}^* óptimos, se utiliza la red ya entrenada para realizar los respectivos pronósticos.

Con la finalidad de analizar el comportamiento de los modelos ante el ingreso de nueva información se utiliza la técnica de pronóstico recurrente conocida como “rolling” en la cual se parte de un grupo de parámetros estimados a partir de un conjunto de datos disponibles hasta el tiempo t ($\bar{\varphi}_{1:t}^J$), con estos parámetros estimados se realiza un pronóstico para $h = 12$. Posteriormente se reestiman estos parámetros con la adición del conjunto de variables correspondientes al periodo $t + 1$ ($\bar{\varphi}_{1:t+1}^J$), y se realiza nuevamente el pronóstico para un horizonte $h = 12$, este proceso se repite i cantidad de veces hasta que existan $t + i - 1$ datos, de manera tal que sea posible comprobar la calidad de pronóstico. Esta metodología se aplica para los tres modelos trabajados, sin embargo para el modelo de redes neuronales se sigue la metodología utilizada por Jalil y Misas (2007) en la cual los valores w_{ij}^* obtenidos del entrenamiento del periodo t son utilizados como pesos “semilla” en la fase de entrenamiento del algoritmo back-propagation del modelo que toma en cuenta la información hasta $t + 1$, lo mismo con $t + 2$, $t + 3$, hasta $t + i - 1$.

1.3. Datos y tratamiento.

Las variables que se tomaron en cuenta para el diseño del modelo VAR y RNA fueron seleccionadas en base a la revisión de los trabajos llevados a cabo por Guerrero (2017), Carrera (2016) y Moran (2014). Como base de la explicación al fenómeno inflacionario en el Ecuador se toma principalmente en consideración el enfoque clásico propuesto por Friedman⁴⁷, por lo que se presta especial atención a las variables de carácter monetario. Las variables con las que se han desarrollado los modelos son:

⁴⁷ El enfoque clásico nos indica que la inflación es un fenómeno monetario, este enfoque es el más aceptado entre los economistas a nivel mundial.

- **Variable de interés:** La variable de interés en el presente estudio es el IPC del Ecuador comprendido entre el periodo 2000-2017, sin embargo ante la presencia de ruido excesivo en la serie provocado por la adopción del dólar como moneda nacional motivó a que se tome en consideración la misma solo desde el año 2004, periodo a partir del cual los cambios inflacionarios empiezan a estabilizarse (Erraez, 2005)⁴⁸.
- **Variables Explicativas:** Como se mencionó anteriormente en este apartado así como en la revisión de literatura correspondiente, para seleccionar las variables explicativas se recopilieron las variables que mayor significatividad tuvieron en otros estudios empíricos realizado para el Ecuador, así como sustento teórico dado por las principales escuelas económicas que han abordado los determinantes de la inflación (escuela clásica, neoclásica y keynesiana). Las variables consideradas son: índice de tipo de cambio real de Colombia (ITCR), tipo de cambio real de Perú (TCR), índices de precios⁴⁹, Ingresos petroleros y no petroleros, gasto corriente del sector público (GCSP), gasto de capital del sector público (GCPSP), salarios, tasa de interés activa referencial productivo corporativo (TIAR), índice de desarrollo de la actividad económica coyuntural (IDEAC), crédito al sector privado (CSP) y los agregados monetarios M1 y M2⁵⁰.

Tal como se puede apreciar en el grafico anexo 4, todas las variables con las que se trabaja se encuentran medidas en escalas diferentes, lo que puede provocar un mal desempeño del modelo RNA, a más de eso muchas ellas presentan un componente de tendencia, lo que se traduce en que las variables

⁴⁸ Los periodos excesivamente ruidosos en la serie del IPC se puede apreciar en la sección Anexos B, apartado Anexo B.1. En las respectivas graficas se puede analizar las tendencias en la serie total y por periodos, donde se precia que una vez apartados los datos correspondientes al periodo 2000-2003 el modelo se ajusta de mejor manera a una función lineal (grafico anexo 3), de la misma manera se observa el aplanamiento de la serie a partir del segundo semestre del año 2016 (periodo desde el que se toma la muestra para la evaluación de pronósticos), razón por la cual se puede esperar problemas en el desenvolvimiento de los modelo para pronósticos mayores que $h = 12$.

⁴⁹ Índices tales como el precio de productos primarios (PCP), índice de precio del petróleo (POILAPSP), índice de precio de la tonelada métrica de banano (PBANSOP), índice de precios de insumos industriales (PINDU). Se seleccionaron estos índices ya que capturan en gran la producción exportada por el ecuador.

⁵⁰ Los gráficos correspondientes a cada variable se aprecian en la sección Anexos, subsección Anexo B.2.

son claramente no estacionarias⁵¹. Sin embargo este problema se solucionó tras diferenciar una vez las variables⁵². Tras la aplicación del test Dickey – Fuller sobre las variables diferenciadas se encontró que estas pasaban a ser variables estacionarias, por lo que este es el orden de integración seleccionado para construir el modelo ARIMA y VAR.

Dado que el objetivo principal del modelo VAR en esta investigación es evaluar el desempeño que este tienen en el pronóstico, el significativo número de variables con las que se está trabajando y la alta correlación que existe entre varias de ellas, genera una menor confianza en el modelo, razón por la cual se probó una reducción de dimensiones a través del análisis de componentes principales.

El análisis de componentes principales es una técnica estadística que consiste en reducir un conjunto de y variables correlacionadas a otro conjunto de z variables llamadas componentes principales, mismos que se encuentran incorrelacionados entre sí. Las variables agrupadas en cada componente son tomadas en función de la variabilidad que estas explican del total de la muestra⁵³, así lo que se busca es obtener $z < y$ combinaciones lineales⁵⁴.

Una práctica habitual previo al análisis de componentes principales es la de tipificar la series⁵⁵ debido a que las diferentes escalas de medida de las series utilizadas pueden influir en los resultados del análisis⁵⁶, también se requiere que las variables presenten un correlación significativa en valor absoluto⁵⁷, caso contrario la reducción de dimensiones pierde sentido.

⁵¹ Para corroborar esto se realizó un test de Dickey – Fuller, mismo que mostro que todas las variables son $I(1)$.

⁵² El grafico de las variables diferenciadas se aprecia en la sección Anexos B, subsección Anexo B.3

⁵³ Para cada componente se eligen variables que representen gran proporción de la variabilidad de la muestra ya que se busca que cada componente tome la mayor información posible.

⁵⁴ El desarrollo matemático del análisis de componentes principales se encuentra detalladamente en el capítulo 3 del libro “*Wiley series in probability and mathematical; applied probability and statistics*”, Jackson (1991).

⁵⁵ Normalización estadística.

⁵⁶ La gráfica de las series normalizadas se aprecia en la sección anexos B, subsección Anexos B.3, Gráfico anexo 6 y 7.

⁵⁷ El test de Bartlett resulta útil para esta función, dicho test, así como otros gráficos relacionados al análisis de componentes principales se aprecia en la sección anexos B, subsección Anexos B.4.

El tratamiento que se le dio a las variables previo al diseño del modelo de RNA fue el de tipificar la serie de manera tal que estas se ajusten de mejor forma a las funciones de activación utilizadas. Así se procedió a realizar la estandarización estadística de las series para la RNA cuyas funciones de activación son lineales, la delimitación de las series a un rango comprendido entre 0 y 1 para la RNA cuyas funciones de activación son sigmoideas logarítmicas y la delimitación de las series a un rango comprendido entre -1 y 1 para la RNA cuyas funciones de activación son tangenciales hiperbólicas⁵⁸.

2. RESULTADOS Y DISCUSIÓN

Tras el análisis de la serie del IPC se observa un proceso desinflacionario en la misma que resulta más pronunciado en el periodo del que se seleccionó los datos para evaluar la calidad de pronóstico, por lo que estas tendencias no formaron parte de la estimación de los parámetros y por lo tanto no pudieron ser capturadas por ninguno de los modelos para el pronóstico del horizonte $h = 18$. Para el pronóstico de horizonte $h = 12$ los modelos muestran un mejor desempeño, aunque los modelos continúan sobre estimando el pronóstico del IPC.

En cuanto a la familia de modelos ARIMA, fueron 4 los modelos que pasaron a la fase de pronóstico tras cumplir con todos los supuestos necesarios en cada una de las etapas de la metodología propuesta por B-J⁵⁹. Ante la presencia de un fuerte componente estacional⁶⁰ los esquemas SARIMA fueron los que mejores resultados arrojaron, tomando en consideración dicho componente ya sea de forma auto regresiva (SAR) o de forma media móvil (SMA).

Los esquemas SARIMA seleccionados para la fase de pronóstico fueron un modelo SARIMA (2,1,1), SARIMA (3,1,1), SARIMA (1,2,1) y SARIMA (1,1,1), de

⁵⁸ Los gráficos de las series tras cada transformación se aprecian en la sección Anexos B, subsección Anexos B.3.

⁵⁹ Los 4 modelos finales fueron seleccionados en base al criterio de Akaike. Las salidas de Eviews, así como las respectivas pruebas de validación de cada modelo se aprecia en la sección Anexos C, subsección Anexos C.1.

⁶⁰ El grafico anexo 16 se observa como el PAC y el AC muestran un fuerte componente estacional, el correlograma corresponde al primer modelo exploratorio el cual pese a que mantiene los errores normalmente distribuidos, estos no están incorrelacionados, es decir no son ruido blanco.

los cuales el modelo SARIMA (1,1,1) fue el que mejores resultados de pronóstico ofreció en base a todos los criterios de evaluación tanto en el corto como en el mediano plazo, tal como se aprecia en la tabla 1.

Tabla 1: Resultados de los modelos ARIMA.

Modelo	ECM	RECM	MPE	MAPE	σ_{MAPE}	U Theil
Corto Plazo						
ARISMA (2,1,1),	3.084152	1.756175	1.4187%	1.4187%	0.00778%	0.008269
ARISMA (3,1,1),	1.681346	1.296667	0.8850%	0.9328%	0.0065%	0.006121
SARIMA (1,2,1)	1.936173	1.391464	1.0316%	1.0316%	0.0069%	0.006564
SARIMA (1,1,1)	0.330690	0.575056	0.3292%	0.4564%	0.0009%	0.002722
Mediano Plazo						
ARISMA (2,1,1),	8.902205	2.983656	2.5006%	2.5006%	0.02046%	0.013984
ARISMA (3,1,1),	5.112455	2.261074	1.7958%	1.3025%	0.0015%	0.010634
SARIMA (1,2,1)	6.193269	2.488628	2.0069%	2.0069%	0.0187%	0.011692
SARIMA (1,1,1)	0.907173	0.952456	0.7513%	0.7513%	0.0027%	0.004503

Elaboración Propia.

Al observar la tabla 1 se puede apreciar claramente a través de todos los indicadores estadísticos como el modelo SARIMA (1,1,1) supera ampliamente en el pronóstico a los otros modelos en los dos horizontes de tiempo, de igual manera al observar la σ_{MAPE} se observa que este modelo es el que menor desviación tiene en sus pronósticos. Sin embargo se debe destacar que tanto para los pronósticos de corto como de mediano plazo, todos los esquemas SARIMA tienden a sobre estimar el pronóstico de la serie (como se observa en el gráfico 6), como ya se mencionó anteriormente, esto es debido al periodo desinflacionario que la serie presenta en los horizontes de pronóstico establecidos⁶¹.

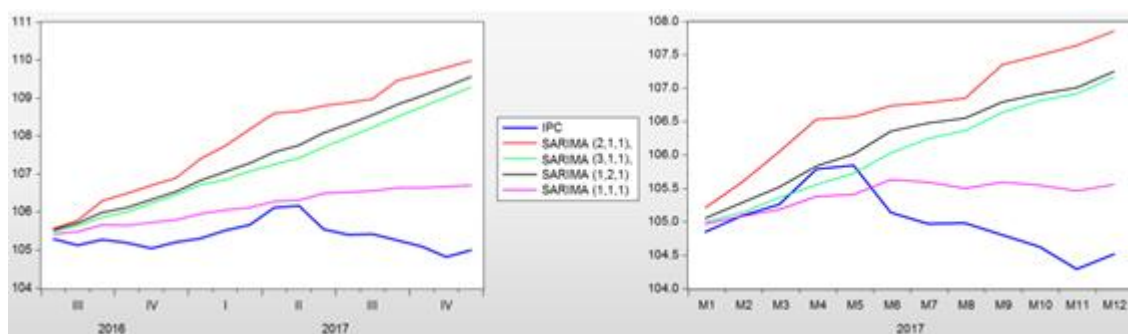


Gráfico 6: Pronóstico de modelos SARIMA mediano y corto plazo
Elaboración Propia.

⁶¹ Como se aprecia en el gráfico 6, para el pronóstico de corto plazo el modelo ya ha detectado el aplanamiento de la serie, por lo que el pronóstico se ajusta de mejor manera al valor original.

Los resultados obtenidos en esta investigación difieren de los obtenidos por Erráez (2006) donde ante la falta de componentes fuertemente estacionales el esquema que mejor modela y pronostica la inflación es el esquema ARIMA (2,1,2), sin embargo este resultado es esperado debido a dos factores clave, el primero es el tamaño de la muestra, mismo que permite detectar de mejor manera patrones anuales de la serie, el segundo es el comportamiento de la serie en cada periodo de tiempo. Erráez lleva a cabo su investigación para el periodo inmediato post dolarización (2000-2004), mientras que para esta investigación se prescindió de esos datos por el exceso de ruido que estos aportaban a los modelos usados.

Para la selección del mejor esquema VAR para el pronóstico de la serie, se consideraron al igual que en el caso de los esquemas ARIMA, aquellos cuyo criterio informacional de Akaike es más reducido. Es importante destacar el hecho de que solamente los modelos con un número alto de rezagos (por sobre los 14 rezagos⁶²) presentan residuos ruido blanco, algo que ya se esperaba tras la selección del modelo SARIMA, el cual muestra la presencia de fuertes componentes estacionales.

El hecho de que los residuos sean ruido blanco tras utilizar un elevado número de rezagos y que la base de datos disponible se limite solamente a 168 observaciones, motivó a que el proceso de la selección de las variables sea mucho más minucioso. Debido a que la cantidad de parámetros a estimar es muy alta para el número de observaciones disponibles en el modelo exploratorio, no se pudo comprobar cuáles son las variables que más influyen en el comportamiento del IPC, razón por la cual para reducir las mismas se procedió de tres maneras diferentes⁶³.

⁶² Las gráficas de los residuos se pueden apreciar en la sección anexos C

⁶³ Como se puede apreciar en la sección anexos C, subsección anexos C.2, el modelo exploratorio presenta un bajo R^2 y los residuos no son ruido blanco, sin embargo este modelo es importante ya que nos permite tener una referencia de la calidad de pronóstico para llevar a cabo la reducción de las variables, razón por la cual es tomado en consideración.

La primera reducción se lleva a cabo a través de los test de significancia individual⁶⁴, seleccionando así solamente las variables cuyo estadístico t relacionado a su parámetro sea más alto, dichas variables fueron los ingresos no petroleros, los salarios, los gastos de capital del sector público y la oferta monetaria.

La segunda estrategia de reducción fue la de tomar las variables que han mostrado mayor relevancia en los trabajos llevados a cabo por Guerrero (2017), Carrera (2016) y Erráez (2005), dichas variables son el tipo de cambio real de Colombia y Perú, tasa de interés activa referencial, precio de los commodities y los salarios.

Estas reducciones fueron realizadas con miras a conseguir un modelo VAR que sirva de ayuda para fines de política económica, sin embargo como se puede apreciar en la subsección anexos C.2, el hecho de incluir un mayor número de rezagos para obtener residuos ruido blanco perjudica seriamente los pronósticos realizados por los modelos, a más de que la raíz inversa del polinomio característico AR esta sobre el círculo unitario (por lo que podría ser un proceso no estacionario), razón por la cual se sacrifica esta propiedad a pos de mejorar los pronósticos y asegurar que estos son válidos en diferentes periodos de tiempo.

La tercera estrategia de reducción de componentes fue la de previo al modelo VAR, realizar un análisis de componentes principales de tal manera que sean las transformaciones lineales dadas por estos componentes junto con la serie del IPC, las variables a utilizar para estimar el modelo VAR.

Cada una de las estrategias llevadas a cabo tuvo diferentes resultados en periodos diferentes de tiempo, siendo el modelo VAR donde se utilizan los componentes principales el más efectivo para llevar a cabo los pronósticos tanto en el corto como en el mediano plazo, siendo este último horizonte de pronóstico en el cual el análisis de componentes principales mostro un

⁶⁴ Se debe recalcar que esto solo con fines exploratorios de pronóstico, ya que como se mencionó anteriormente, ante correlación serial los test de significancia tanto individual como colectiva carecen de confianza.

incremento significativo en el desempeño del modelo, tal como se aprecia en la tabla 2.

En cuanto a los modelos VAR destinados para fines de política, las estrategias de reducción también dieron resultados satisfactorios, siendo estos más notorios en el pronóstico de mediano plazo, horizonte en el cual el modelo VAR donde se utilizan solo las variables explicativas, mejora significativamente la calidad de pronóstico en comparación a los otros modelos VAR. Para el corto plazo se observa que la mejor estrategia fue la de utilizar las variables dadas por la literatura, sin embargo estos resultados no difieren mucho del modelo exploratorio donde se utilizan todas las variables en logaritmos.

Tabla 2: Resultados de los modelos VAR⁶⁵

Modelo	ECM	RECM	MPE	MAPE	σ_{MAPE}	U Theil
Corto Plazo						
VAR 1	2.564664	1.601457	1.1528%	1.1528%	0.00994%	0.007550
VAR 2	6.084248	2.466627	1.9784%	1.9784%	0.0159%	0.011581
VAR 3	5.357771	2.314686	1.8643%	2.2845%	0.0049%	0.011825
VAR 4	7.152412	2.674399	1.8643%	1.8643%	0.0137%	0.010874
VAR 5	2.533487	1.591693	1.0734%	1.1049%	0.0107%	0.007507
VAR 6	2.513194	1.585305	1.1803%	1.1803%	0.0088%	0.007473
Mediano Plazo						
VAR 1	13.54679	3.680596	2.8310%	2.8310%	0.04845%	0.017221
VAR 2	9.640707	3.104948	2.6038%	2.6038%	0.0224%	0.014545
VAR 3	9.005316	3.000885	2.5022%	2.5022%	0.0219%	0.014064
VAR 4	11.95859	3.45812	2.8202%	2.8202%	0.0341%	0.016181
VAR 5	9.882945	3.143715	2.6478%	2.6478%	0.0223%	0.014723
VAR 6	4.757118	2.181081	1.7873%	1.7873%	0.0139%	0.010258

Elaboración Propia.

Al comparar los resultados obtenidos con los esquemas VAR se puede apreciar un desempeño significativamente peor que los esquemas SARIMA, ya que como se aprecia en el grafico 7 todos estos esquemas VAR tienen problemas para capturar el periodo desinflacionario que presenta la serie, por lo que todos los pronósticos son sobre estimados.

⁶⁵ VAR 1: modelo exploratorio con todas las variables en logaritmo y 1 rezago, VAR 2: modelo exploratorio con todas las variables y 1 rezago, VAR 3: modelo con variables explicativas, VAR 4: modelo con variables explicativas en logaritmo, VAR 5: modelo con variable de la revisión de la literatura, VAR 6: modelo componentes principales.

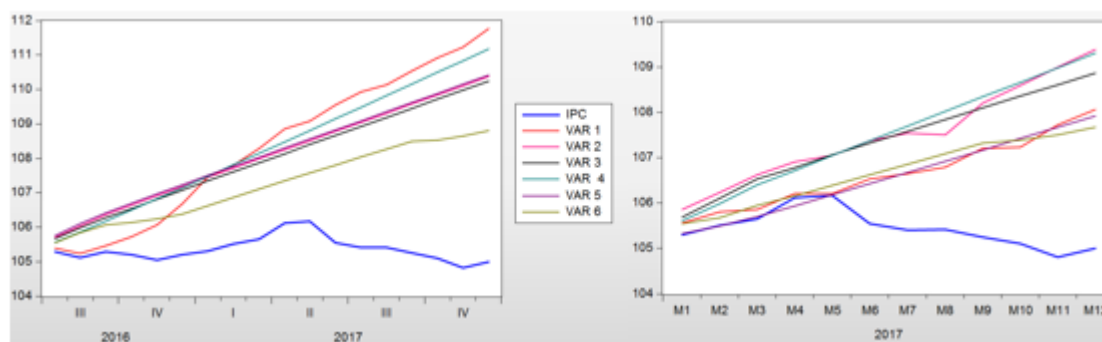


Gráfico 7: Pronóstico de modelos VAR mediano y corto plazo
Elaboración Propia.

Los resultados obtenidos con el desarrollo del modelo VAR son muy similares a los obtenidos por Guerrero (2017) donde también requiere un gran número de rezagos para acercarse a obtener residuos ruido blanco y poder establecer causalidad entre variables⁶⁶. De la misma manera Chiquito (2014) tiene que recurrir a un alto número de rezagos (10 rezagos) para conseguir que su modelo tenga perturbaciones heterocedasticas, así como la inclusión de variables dummy que capturan el comportamiento estacional de la economía ecuatoriana⁶⁷.

Una razón importante para no utilizar un alto número de rezagos y pasar de la inclusión de variables dummy es que estos provocaron una especie de sobre ajuste del modelo, lo que no permitió una correcta generalización del mismo y perjudicó seriamente la calidad de los pronósticos. Se prestó especial atención a este problema para el modelo VAR exploratorio, ya que los pronósticos realizados por este, serán los que sirvan para alimentar al modelo RNA.

Con respecto al modelo de RNA se puede optar por 2 tipos de red, la primera es seguir una topología de red tipo ARMA, el inconveniente de utilizar esta topología de red es que los resultados obtenidos por el modelo rara vez son mejores que los modelos ARIMA cuando la serie se presenta estable y sin mucho ruido⁶⁸. La segunda topología es de tipo modelo multivariante en el que

⁶⁶ El objetivo central de la investigación de Guerrero es encontrar los factores determinantes de la inflación.

⁶⁷ En la presente investigación se encontró que la inclusión de variables dummy efectivamente mejoraba el comportamiento de los residuos.

⁶⁸ Es importante destacar que el desempeño del modelo SARIMA (1,1,1) utilizado para el pronóstico del IPC es bastante ajustado a la serie original, razón por la que intentar adoptar una topología de red de este tipo no aporta mucho a la investigación.

se busca el pronóstico de una serie en base de un conjunto de variables explicativas, que es la topología por la que se optó en este trabajo.

El problema de utilizar una topología multivariante para llevar a cabo el pronóstico de series macroeconómicas es que en la mayoría de casos no está disponible la información ni de la variable de interés, ni de las variables explicativas⁶⁹, por lo que estas últimas primero deben ser pronosticadas y con estas realizar el pronóstico de la serie de interés.

En cuanto a la estrategia utilizada para realizar el pronóstico de las variables explicativas del IPC se utilizó el modelo VAR exploratorio (VAR 1) el cual pese a no ser el modelo que mejores resultados ofreció en el pronóstico del total de modelos VAR probados, para fines prácticos es el que mejor cumple con los requerimientos de los modelos RNA⁷⁰.

En los modelos de RNA a diferencia de los modelos VAR mientras mayor sea el número de entradas que alimenten la red, mayor es la capacidad de generalización de la misma, es decir que para un mejor funcionamiento de la red se requiere el pronóstico de todas las variables explicativas que se tomaron en consideración en un principio⁷¹.

Una vez obtenido el conjunto de variables explicativas pronosticadas, se procedió a seleccionar las diferentes arquitecturas de red, siendo las topologías de tipo perceptrón multicapa y red recurrente las seleccionadas para realizar el pronóstico⁷². Para el caso de las neuronas, se probó los resultados obtenidos con cada función de activación y sus respectivas transformaciones de variables, para los pesos sinápticos iniciales se eligió un valor al azar al igual que los valores umbrales, mismos que tras las iteraciones en el proceso de

⁶⁹ Por lo general los Bancos Centrales y otras instituciones como el INEC brindan la información en conjunto a través de boletines publicados con periodicidad mensual, trimestral, etc.

⁷⁰ El pronóstico conjunto de las variables realizado por el modelo VAR resulta muy práctico debido al ahorro en el tiempo y la eficiencia en los resultados obtenidos

⁷¹ Los resultados obtenidos con la utilización de las series originales y de componentes principales, así como los resultados obtenidos por las demás redes se pueden apreciar en la sección Anexos C.3.

⁷² Las diferentes arquitecturas de red usadas se aprecia en la sección Anexos C, subsección Anexos C.3, apartado Modelo RNA tipo recurrente con función de activación tangente hiperbólica y sigmoidea logarítmica.

entrenamiento convergen a sus valores óptimos, en cuanto a la regla de propagación se optó por un modelo aditivo.

Las arquitecturas de red cuyas funciones de activación fueron lineales dieron mejores resultados que aquellas redes que utilizaron funciones de activación sigmoidea logarítmica y tangencial hiperbólica⁷³. Estos resultados fueron esperados tras el análisis del comportamiento de la serie y su buen ajuste a una función lineal, sin embargo no se esperaba que la diferencia sea tan significativa, ya que las mejores RECM de los modelos RNA_{SL} y RNA_{TH} en el mediano plazo fueron de 4.68 y 4.76 respectivamente, a diferencia de la mejor RECM obtenido en la red cuyas funciones de activación fueron lineales, el cual fue de 1.81.

Una vez seleccionada la función de activación, se procedió a elegir el número de capas ocultas, el número de neuronas de cada capa y la manera en la que las mismas se encuentran interconectadas. Como punto de partida se tomaron en consideración los diferentes estudios llevados a cabo, mismos que señalan que una o máximo dos capas ocultas son ideales para las RNA de series de tiempo macroeconómicas, también se tomó en consideración el hecho de que muchas veces un número excesivo de capas ocultas, así como de neuronas ocultas, producen un sobre ajuste del modelo.

La estrategia llevada a cabo para la selección de capas y neuronas ocultas fue un proceso de prueba y error, empezando con una capa oculta en un principio e incrementando el número de capas sucesivamente, cada arquitectura fue entrenada y se realizaron sus respectivos pronósticos, los cuales se evaluaron bajo los criterios antes especificados. Una vez escogido el número de capas se procedió de igual manera para seleccionar el número de neuronas en cada capa.

En cuanto a la etapa de entrenamiento se seleccionó un conjunto de 30 variables (aproximadamente el 20% de la muestra) a las que se sometieron a un proceso de entrenamiento back-propagation con detención temprana (cuyos

⁷³ Los resultados de pronóstico, así como las arquitecturas de red utilizadas se aprecia en la sección anexos C, subsección Anexos C.3.

criterios fue el número de iteraciones o el error mínimo establecido⁷⁴). Por último, para la selección del ratio de aprendizaje en cada época o iteración, se tuvo que recurrir a un procedimiento tipo prueba y error y testear aquel η que mejor resultado arrojó, siendo este valor de $\eta = 0.01$. Estos parámetros fueron los seleccionados para todos los modelos.

Del total de los modelos probados, se determinó que el que mejores resultados en el pronóstico de mediano plazo dio fue la red tipo recurrente con una capa oculta, 8 neuronas ocultas y una función de activación lineal en todas las capas. La fase de entrenamiento se detuvo en un total de 13661 iteraciones. Este modelo fue el que se seleccionó también para el pronóstico de corto plazo realizando los ajustes que la modificación del tamaño de muestra requirió⁷⁵.

Como se puede apreciar en la tabla 3 los resultados de los modelos RNA en el pronóstico de corto plazo son apenas peores que los pronósticos realizados por el modelo VAR, en parte debido a que los parámetros del modelo de largo plazo fueron los mismos que se utilizaron en el modelo de corto plazo. Donde se aprecia realmente una mejora significativa es en el pronóstico del mediano plazo, donde el modelo de RNA tiene una significativa mejora en comparación a todos los modelos VAR testeados.

Tabla 3: Resultado modelo RNA de corto y mediano plazo

Modelo	ECM	RECM	MPE	MAPE	σ_{MAPE}	U Theil
Corto Plazo						
RNA CP	2.770687	1.664538	1.2531%	1.2531%	0.00939%	0.007844
Mediano Plazo						
RNA MP	3.259163	1.805315	1.4341%	1.4341%	0.0105%	0.008506

Elaboración Propia.

Pese a que los modelos RNA consiguen mejorar los pronósticos de los modelos VAR en el mediano plazo y mantener un pronóstico muy similar en el corto plazo, es importante destacar que estos también presentan dificultades para capturar el periodo desinflacionario de la serie y parecen continuar la tendencia haciendo caso omiso del mismo (como se aprecia en el gráfico 8), por lo que el pronóstico ira empeorando conforme el horizonte h se incrementa.

⁷⁴ El número máximo de iteraciones se estableció en 100000 y el error mínimo se estableció en 0.05.

⁷⁵ Se procedió a reestimar el modelo con un nuevo tamaño de muestra y a reentrenar la red, todo bajo la misma arquitectura de red y con los mismos parámetros de la red de mediano plazo.

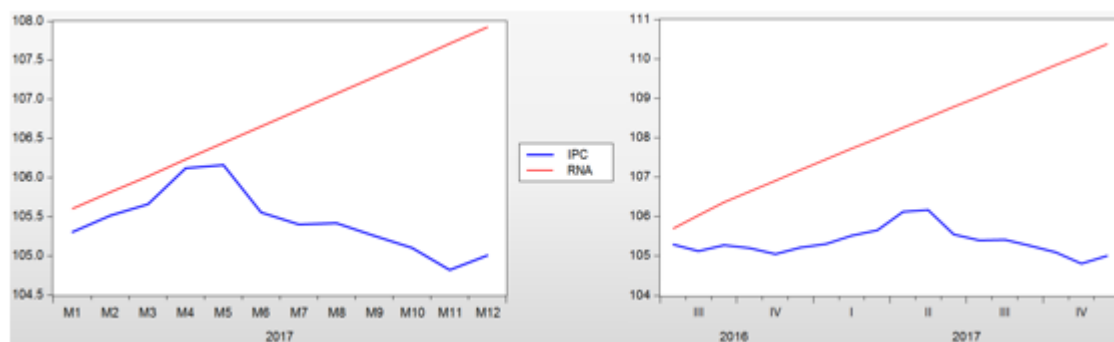


Gráfico 8: Pronóstico modelo RNA corto y mediano plazo
Elaboración Propia.

Los resultados obtenidos en cuanto a los modelos de RNA son bastante similares a los obtenidos en el estudio del Banco Central Europeo (2004) donde los modelos de RNA que mejores resultados ofrecen son aquellos con funciones de activación lineales, dichos modelos consiguen un buen ajuste en los pronósticos de aquellos índices de precios cuyos datos no siguen una tendencia lineal, como los precios en el sector de los servicios.

Salazar (2009) obtiene resultados similares a los encontrados en este estudio, salvo la excepción de que en su trabajo los modelos de RNA consiguen mejorar el pronóstico en todos los horizontes de tiempo, lo que según el autor, se debe a la existencia de no linealidades en la serie trabajada, escenario ante el cual los modelos RNA se desenvuelven de mejor manera que los modelos tradicionales.

Por ultimo Santana (2006) en su estudio también consigue mejorar los pronósticos del IPC colombiano a través de modelos de RNA, sin embargo a diferencia de este estudio, Santana consigue modelos cuyos pronósticos son mejores que los modelos SARIMA y suavizamiento exponencial. Santana al igual que en el presente estudio, encuentra evidencia de la mejora en los pronósticos que tiene el hecho de combinar diferentes técnicas de modelado.

Tal como se puede apreciar en la tabla 4, el modelo SARIMA (1, 1, 1) es el que mejores resultados ofrece tanto en el corto como en el mediano plazo cuando los objetivos son solamente el pronóstico de la serie. Sin embargo cuando los objetivos son de política económica y es necesario recurrir a un modelo multivariante como el modelo VAR, la aplicación de modelos RNA posterior a la

modelación VAR puede mejorar los resultados de pronóstico, sobre todo en horizontes más largos de tiempo.

Tabla 4: Resultados del pronóstico de los diferentes modelos.

Modelo	ECM	RECM	MPE	MAPE	σ_{MAPE}	U Thai
Corto Plazo						
SARIMA (1,1,1)	0.330690	0.575056	0.3292%	0.4564%	0.0009%	0.002722
VAR 5	2.533487	1.591693	1.0734%	1.1049%	0.0107%	0.007507
RNA CP	2.770687	1.664538	1.2531%	1.2531%	0.00939%	0.007844
Mediano Plazo						
SARIMA (1,1,1)	0.907173	0.952456	0.7513%	0.7513%	0.0027%	0.004503
VAR 3	9.005316	3.000885	2.5022%	2.5022%	0.0219%	0.014064
RNA MP	3.259163	1.805315	1.4341%	1.4341%	0.0105%	0.008506

Elaboración Propia.

Por último, con los modelos seleccionados de cada familia se probó la estabilidad que estos presentan ante el ingreso de nueva información a través de la metodología de pronóstico recurrente, para llevar a cabo esta metodología se procedió a cortar la serie 18 periodos antes del último dato disponible, con este conjunto de datos se estiman los modelos y se realizan los pronósticos $h = 12$ periodos adelante. Para la segunda etapa del pronóstico recurrente, se ingresa el valor un periodo t adicional, se reestiman los parámetros y se realiza el pronóstico para el subsecuente horizonte $h = 12$. Este procedimiento se repite hasta tener $T - 1$ observaciones ya que es hasta este periodo donde se puede evaluar la calidad de pronóstico.

En el análisis se realizaron 6 iteraciones, cuyos resultados se pueden apreciar en el gráfico 9. En este gráfico se observa como el modelo SARIMA es el que mayor estabilidad presenta, sin embargo se observa que este tiene una ligera tendencia a incrementar la RECM en cada iteración realizada. En cuanto a los modelos VAR y RNA se observa como tienen un comportamiento un tanto más ruidoso, observando que el modelo VAR tiende también a incrementar la RECM con cada iteración, mientras el modelo RNA no tiene una tendencia clara. Las implicancias de los resultados obtenidos en este proceso, sobre todo para los modelos de RNA, es que estos son muy susceptibles a las modificaciones en los parámetros del proceso de entrenamiento y para esta investigación se utilizaron los mismos parámetros para cada iteración, sin considerar posibles cambios en las tendencias ni los ajustes necesarios ante la

variación en el tamaño de la muestra, lo que se puede apreciar en la calidad de pronóstico tras cada iteración.

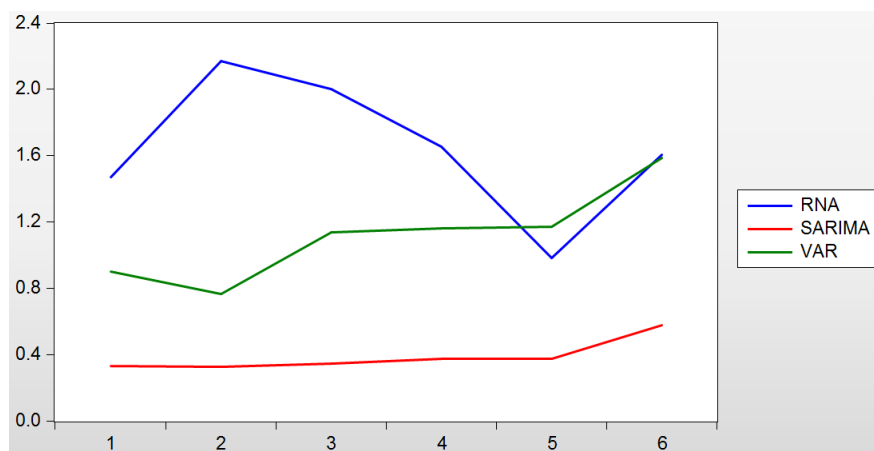


Gráfico 9: RECM diferentes modelos en cada iteración del pronóstico recurrente.
Elaboración Propia.

Es importante destacar que todos los modelos se presentan estables en las 6 iteraciones realizadas, por lo que los resultados obtenidos por los tres modelos pueden ser considerados confiables, al menos para el periodo de tiempo analizado. De la misma manera se debe destacar la indudable superioridad de los esquemas SARIMA tanto en el pronóstico de corto y mediano plazo, así como en la estabilidad que presenta el modelo.

Con respecto a los esquemas multivariantes, es destacable la mejora que los modelos de RNA en conjunto con los modelo VAR tiene para los pronósticos, sobre todo de mediano plazo. Para el caso del corto plazo se considera que la adopción de otras topologías de red, otras técnicas de obtención de las variables explicativas o diferentes tratamientos a las variables pueden mejorar notablemente el pronóstico de la red de CP.

3. CONCLUSIONES Y RECOMENDACIONES.

Como se ha podido observar a lo largo de la investigación, cuando la modelación de la series solamente tienen fines de pronóstico, los esquemas SARIMA son la mejor opción de modelado ya que estos superan ampliamente tanto a los modelos VAR como a los modelos de RNA, esto tanto en la calidad de pronóstico como en la estabilidad del modelo para los dos periodos de tiempo analizados (corto y mediano plazo).

Sin embargo cuando el objetivo de estudio es la obtención de modelos que ayuden a la toma de decisiones de política económica, se observa como el análisis de las variables previo a la modelación, las diferentes estrategias para la selección de las variables y el tratamiento que se le dé a las mismas, puede afectar de manera significativa a los resultados obtenidos, dichas estrategias han mostrado ser más o menos eficaces dependiendo del horizonte de pronóstico evaluado.

Es importante destacar la significativa mejoría en la calidad de pronóstico obtenido por el modelo VAR cuando, previo a la modelación del mismo, se ha realizado un análisis de componentes principales. Dicha mejoría de pronóstico mostro ser más evidente en los horizontes más largos de tiempo, donde el modelo VAR elaborado con los componentes principales supera ampliamente a todos los demás modelos VAR tentativos, tanto en la calidad de pronóstico como en la estabilidad del modelo. Para periodos cortos de tiempo el resultado es el mismo, aunque la diferencia entre los diferentes modelos VAR es menos marcada que en el mediano plazo.

En cuanto a las reducciones de variables llevadas a cabo para la obtención de modelos VAR con fines de política económica, se observa que tanto la selección de variables a través de la revisión de la literatura, así como la selección de variables en base a su nivel de significancia dieron buenos resultados, aunque estos fueron mejores en horizontes más largos de pronóstico.

En el corto plazo, la selección de variables a través de la revisión de la literatura fue la que mejores resultados dio, sin embargo los resultados obtenidos son muy similares a los que se obtuvieron al utilizar todas las variables en logaritmos. Para el mediano plazo la estrategia que mejores resultados obtuvo fue la selección de las variables según su nivel de significancia, sin embargo se debe destacar que los resultados tampoco difieren en gran medida de los demás modelos VAR con los que se trabajó.

Es importante mencionar que todos los modelos VAR con los que se ha trabajado no tienen errores ruido blanco, esto debido a que para ello se

requiere un elevado número de rezagos o la introducción de variables dummy que capturen la estacionalidad de las series. Las dos medidas que se tomaron para intentar obtener residuos bien comportados afectaron seriamente la calidad de pronóstico de los modelos, por un lado, un elevado número de rezagos posiblemente hace que el proceso sea no estacionario, por otro lado, la introducción de variables dummy genera una especie de sobre ajuste en el modelo, lo cual impide una correcta generalización del mismo.

En cuanto a los modelos de RNA, en primer lugar se encontró que este funciona muy bien como un complemento del modelo VAR, ya que como se mencionó anteriormente, son los pronósticos de este los que se utilizan como las variables de entrada que alimentan la red para realizar el pronóstico de la serie de interés. Es importante destacar que para el pronóstico del corto plazo, los resultados obtenidos por el modelo VAR y el modelo de RNA son muy similares, sin embargo para el pronóstico a horizontes más largos de tiempo, los modelos de RNA si mostraron ser mucho más eficientes que todos los modelos VAR.

En cuanto a las topologías de red utilizadas, se encontró que la red neuronal de tipo recurrente mejora significativamente la calidad de pronóstico en comparación con la red tipo perceptrón multicapa. También se constató la eficiencia de la utilización de una sola capa oculta con un número reducido de neuronas ocultas, mismas que cuentan con una función de activación lineal.

En cuanto a la etapa de entrenamiento de la RNA, se encontró que la técnica back propagation con detención temprana es la mejor estrategia para evitar el sobreajuste del modelo, imponiendo un número máximo de iteraciones en 100000 y un error máximo de 0.05, con la finalidad de evitar mínimos locales el ratio de aprendizaje $\eta = 0.01$ mostro ser el que mejores resultados ofrece.

Como se aprecia claramente, cuando la modelación tiene fines solamente de pronóstico, el modelo SARIMA (1, 1, 1) es el que mejores resultados ofrece, sin embargo cuando se buscan modelos de soporte para política económica, la combinación de los modelos VAR y de RNA tiene ventajas significativas en comparación con la utilización de cada modelo por separado.

Al no tener modelos VAR con residuos ruido blanco, este se ve muy limitado en cuando a fines de apoyo a política económica, ya que no es posible obtener estadísticos de significancia individual o global confiables, es decir no se puede asegurar la relación existente entre variables, por lo tanto tampoco se puede confiar en las funciones impulso respuesta que se obtengan en este modelo.

Las estrategias llevadas a cabo en esta investigación para solucionar los residuos mal comportados tampoco mostraron ser eficientes, razón por la cual se recomienda trabajar más profundamente en la especificación del modelo. La introducción de variables capaces de capturar las expectativas de la población, el verdadero poder adquisitivo de las personas, el contexto social, económico o político nacional puede tener una mejora significativa en el comportamiento del modelo.

El incremento del tamaño de la muestra también puede ayudar a mejorar los resultados obtenidos por el modelo ya que permitiría, en primer lugar poder utilizar los datos que muestran el periodo desinflacionario de la serie en la etapa de estimación del modelo y así poder obtener mayor precisión en la fase pronóstico, en segundo lugar nos permite utilizar un mayor número de rezagos, lo que puede solucionar el problema en los residuos.

En cuanto a los modelos de RNA, un punto crítico que puede estar afectando el desempeño del mismo es la utilización de los pronósticos de las variables explicativas del modelo VAR, es decir que el modelo de RNA puede estar heredando los problemas que el modelo VAR tiene, razón por la cual el modelo de RNA tampoco capta de buena manera la etapa desinflacionaria que atraviesa el Ecuador. Por esta razón, la utilización de diferentes técnicas para la obtención de las variables explicativas o la mejora del modelo VAR puede significar la mejora del modelo de RNA.

Otra falencia que puede afectar al modelo de RNA, sobre todo al modelo de corto plazo, es que para la obtención de los pesos y las conexiones dentro de la red, se utilizaron los mismos parámetros que en el modelo destinado al pronóstico del mediano plazo, esto podría ocasionar una mala generalización



del modelo, razón por la cual los pronósticos del modelo VAR y de RNA son muy similares en el corto plazo.

Como se puede observar, uno de los factores más importantes para el desempeño de la red es la topología que se adopte en la misma, se debe aclarar que las topologías de red adoptadas en este estudio corresponde a algoritmos diseñados para análisis que no corresponden a las ciencias sociales, por lo que una mayor profundización en este campo, puede llevar a la creación de algoritmos especializados que le permitan a las ciencias sociales contar con redes especializadas en esta área, cuyos resultados podrían ser mejores que los obtenidos con las topologías aquí utilizadas.



4. BIBLIOGRAFÍA

- Box, G., Jenkins, G., & Reinsel, G. (1994). *Time series analysis: forecasting and control*. New Jersey: Prentice-Hall.
- Carrera, C. (Abril de 2016). The inflation – output GAP dynamics in Ecuador: A Descriptive Analysis during Episodes of Positive and Negative Output Gaps. 1970 – 2014. Seul, Corea del Sur: Seul National University.
- Chen, T. (12 de Marzo de 2001). Global exponential stability of delayed Hopfield neural networks. Shangai, China R.P.: Pergamon.
- Echevarría, S. (1975). El pronóstico economico como instrumento de la planificacion empresarial. Alcala, España: Universidad de Alcal, Escuela Superior de Gestión Comercial y Marketing.
- Enders, W. (2009). *Applied Econometric Time Series*. Iowa: John Wiley & Sons.
- Erraez, P. (2005). El proceso inflacionario en el Ecuador: análisis de sus determinantes con modelos ARIMA y vectores auto regresivos. *Banco Central del Ecuador Cuestiones Económicas*, 47-96.
- Fisher, I. (1922). The Purchasing Power of Money, its Determination and Relation to Credit, Interest and Crises. New York, Estados Unidos: Macmillan, new and revised edition.
- Friedman, M. (Marzo de 1970). The Counter-Revolution in Monetary Theory. New York, Estados Unidos: American Economic Review.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. Massachusettes: MIT Press, Cambridge.
- Guerrero, S. (Noviembre de 2017). Determinantes de la inflacion en el Ecuador 2000-2016. Buenos Aires, Argentina: Universidad de Buenos Aires, Facultad de Ciencias Economicas .
- Gujarati, D., & Porter, D. (2009). *Econometria 5th edition*. Mexico: McGRAW-HILL/INTERAMERICANA EDITORES, S.A. DE C.V.
- Guzaman, M. d., & Alba, P. (2008). El modelo VAR y sus principales problemas. *Panorama Economico*, 95-117.
- Hayin, S. (2009). *Neural networks and learning machines* . New Jersey : Pearson education Inc.
- ILO, IMF, OECD, UNECE, Eurostat, & Bank, T. W. (2004). Consumer price index manual: Theory and practice. Ginebra: Organizacion Internacional del Trabajo.
- Instituto Nacional de Estadísticas y Censos. (Enero de 2015). METODOLOGÍA DEL ÍNDICE DE PRECIOS AL CONSUMIDOR (IPC) BASE ANUAL: 2014=100. Quito, Ecuador: Ecuadorencifras.



- Iqbal, M., & Naveed, A. (Enero de 2016). Forecasting Inflation: Autoregressive Integrated Moving Average Model. Vienna, Austria: European Scientific Journal.
- Jackson, E. (1991). Wiley series in probability and mathematical; applied probability and statistics. Estados Unidos : John Wiley & Sons, Inc.
- Jalil, M. A., & Misas, M. (2007). Evaluación de pronósticos del tipo de cambio utilizando redes neuronales y funciones de perdida asimetrica. *Revista colombiana de estadística*, 143-161.
- Keynes, J. (Febrero de 1936). The General Theory of Employment, Interest and Money. Londres, Reino Unido : Palgrave Macmillan.
- Laptev, N., Yosinski, J., Errran, L., & Smyl, S. (2017). Time-series Extreme Event Forecasting with Neural Networks at Uber. *International Conference on Machine Learning*, 1-5.
- McNelis, P., & McAdam, P. (Abril de 2004). Forecasting inflation with thick models an neural networks. Franfut, Alemania: Banco Central Europeo.
- Morán, D. (20 de Noviembre de 2014). Determinantes de la inflacion en Ecuador: Un analisis econometrico utilizando modelos VAR . Mexico , Mexico : Universidad Autonoma Metropolitana.
- Naciones Unidas. (1 de Enero de 2017). Nueva metodologia del indice de precios al consumidor (IPC) del Ecuador (base anual: 2014=100). Santiago, Chile: Naciones Unidas.
- Nakamura, E. (21 de Noviembre de 2004). Inflation forecasting using a neural network. *Economic letters* , 373-378.
- Pindyck, R. S., & Rubinfeld, D. L. (2001). *Econometría: modelos y pronostico*. Mexico: McGRAW-HILL/INTERAMERICANA EDITORES.
- Russell, S., & Norving, P. (2010). *Artifitial intelligence a modern approach, 3rd edition*. New Jersey: Pearson Education Inc.
- Salazar, J. (2009). Evaluacion de pronostico de un red neuronal sobre el PIB en Colombia . *Borradore de Economia* , 1-55.
- Salazar, M. (Enero de 2009). Evaluacion de pronostico de una red nueronal sobre el PIB en Colombia . Bogota , Colombia: Pontificia Universidad Javeriana .
- Santana, C. (Junio de 2006). Predicción de series temporales con redes neuronales: una aplicacion a la inflacion colombiana. *Revista Colombiana de Estadística* , 77-92.
- Sims, C. (1986). Are forecasting models usable for policy analysis. *Federal Reserve Bank of Minneapolis* , 1-16.
- Sims, C. A. (1980). Macroeconomics and Reality . *Econometrica*, 1-48.
- Stock, J. H., & Watson, M. W. (2012). *Introducción a la Econometría 3ra edición*. Madrid: Pearson Education.
- Tostado, S., Ornelas, M., Espinal, A., & Puga, H. (2015). *Implementación de algoritmos de inteligencia artificial para el entrenamiento de redes neuronales de segunda*



generacion. Leon, Mexico: Tecnológico nacional de Mexico, Division de estudio de postgrado e investigacion.

Velazquez, J., & Franco, C. (1 de noviembre de 2010). Nota sobre la prediccion del indice de precios al consumidor usando redes neuronales artificiales. Bogota, Colombia: Universidad Nacional de Colombia.

Veléz, J., & Nieto, P. (2016). Validación de medidas de evaluación para el pronóstico de la tasa de cambio en Colombia. Bogota, Colombia: Colegio de estudios superiores de administracion (CESA).

Wang, L., Zeng, Y., & Chen, T. (2014). Back propagation neural network with adaptive differential evolution algorithm for time series forecasting. *Expert Systems with Applications*, 1-9.

5. ANEXOS

A. Medidas de calidad de pronóstico.

Tabla Anexo 1: Medidas estadísticas de calidad de pronóstico

Medida	Formula	Descripción
Error cuadrático medio (ECM)	$ECM = \frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y}_i)^2$	Medida de que da el promedio del error de pronóstico del periodo h . Permite conocer la incertidumbre alrededor del pronóstico.
Raíz del error cuadrático medio (RECM)	$RECM = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y}_i)^2}$	Informa sobre el tamaño medio del error de pronóstico del periodo h de manera más intuitiva.
Error porcentual medio (MPE)	$MPE = \sum_{i=1}^N \frac{Y_i - \bar{Y}_i}{Y_i}$	Valor que fluctúa entre 0 y 1 que nos permite saber si los pronósticos son en promedio sobre o sub estimados.
Promedio de error porcentual absoluto (MAPE)	$MAPE = \frac{1}{N} \left(\sum_{i=1}^N \frac{ Y_i - \bar{Y}_i }{Y_i} \right) * 100$	Medida del error porcentual promedio de pronóstico, útil para realizar comparaciones entre series de tiempo diferentes.
Desviación estándar de APE ⁷⁶	$\sigma_{MPE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (APE_i - MAPE)^2}$	Medida porcentual del grado de incertidumbre del pronóstico.
Exactitud de pronóstico U1 de Theil	$U_1 = \frac{\sqrt{\sum_{i=1}^N (Y_i - \bar{Y}_i)^2}}{\sqrt{\sum_{i=1}^N Y_i^2} + \sqrt{\sum_{i=1}^N \bar{Y}_i^2}}$	Brinda un índice de la exactitud del pronóstico del modelo.

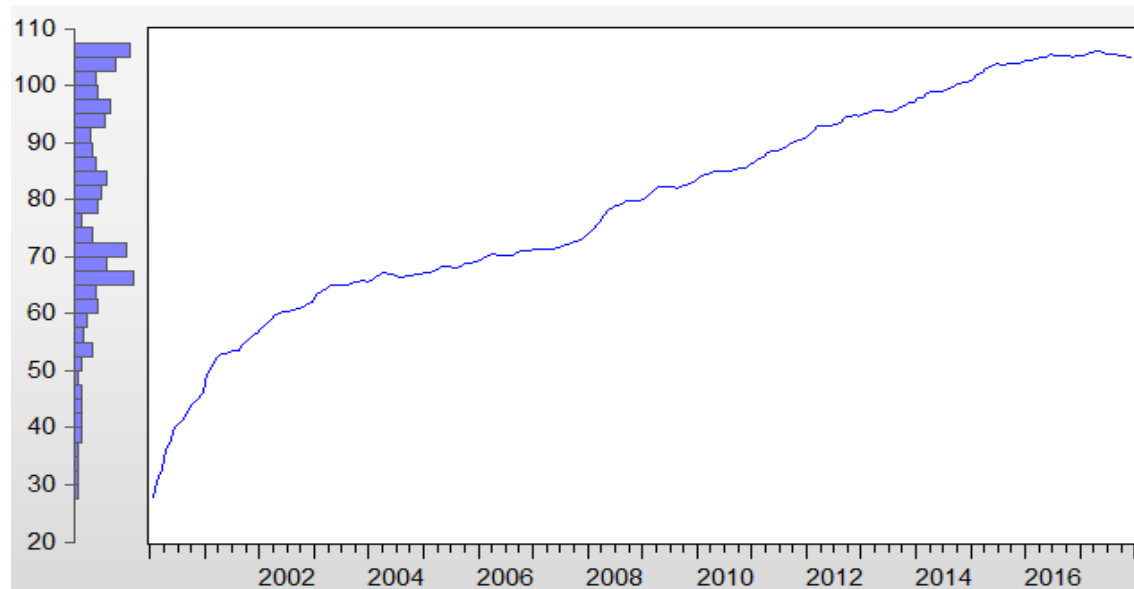
Elaboración Propia.

⁷⁶ $APE_i = \frac{|Y_i - \bar{Y}_i|}{Y_i}$

B. Variables

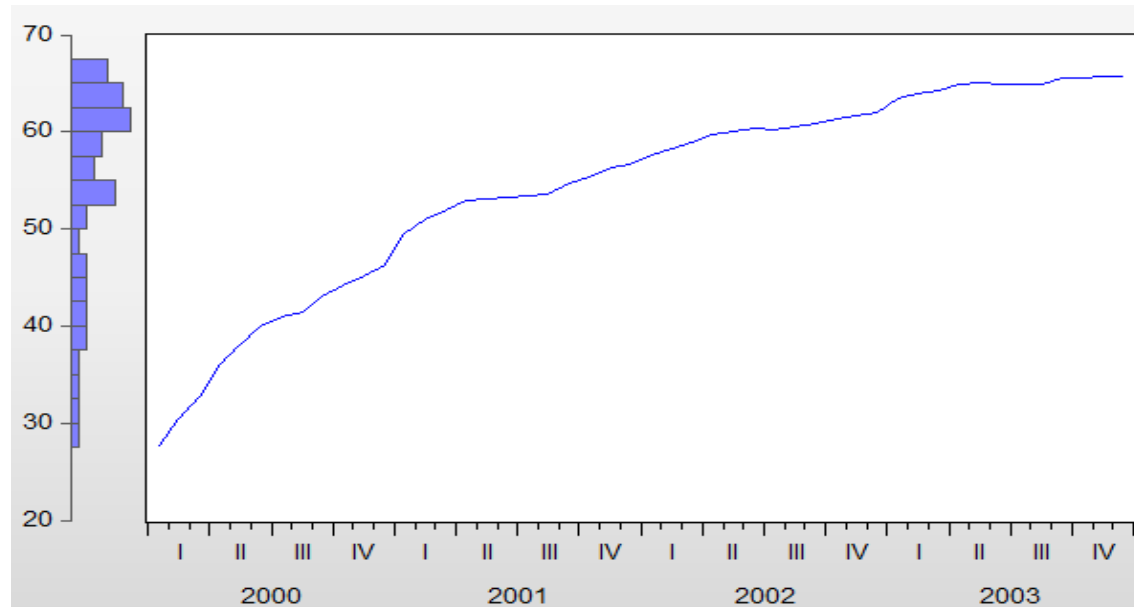
B.1. Variable de interés.

Gráfico Anexo 1: IPC 2000-2017



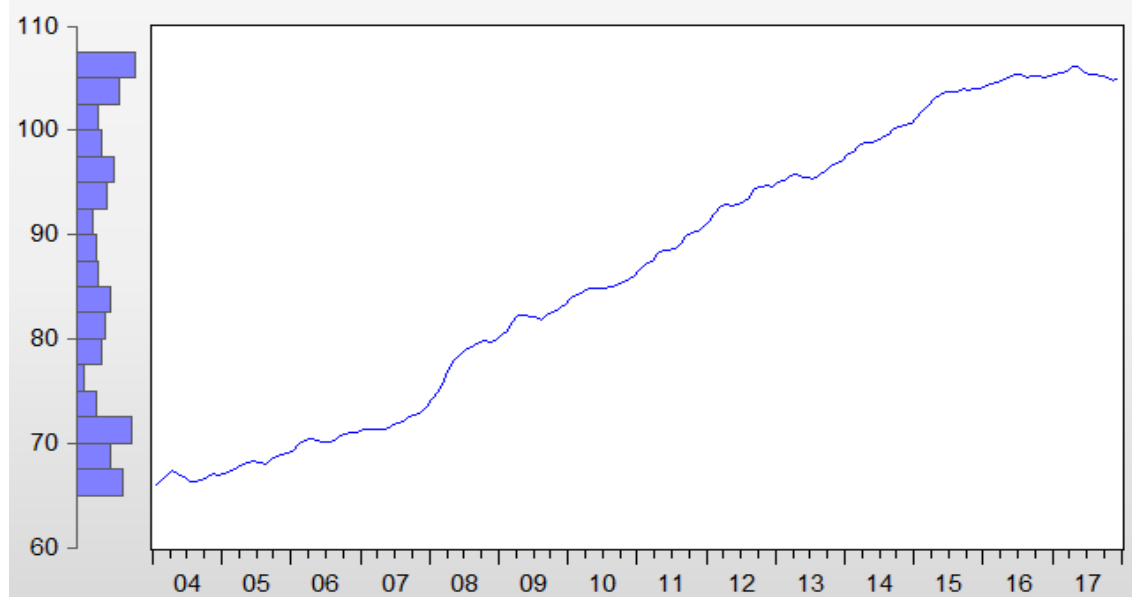
Fuente: Banco Central del Ecuador (BCE)
Elaboración propia.

Gráfico Anexo 2: IPC 2000-2003



Fuente: Banco Central del Ecuador (BCE)
Elaboración propia.

Gráfico Anexo 3: IPC 2004-2017



Fuente: Banco Central del Ecuador (BCE)
Elaboración propia.

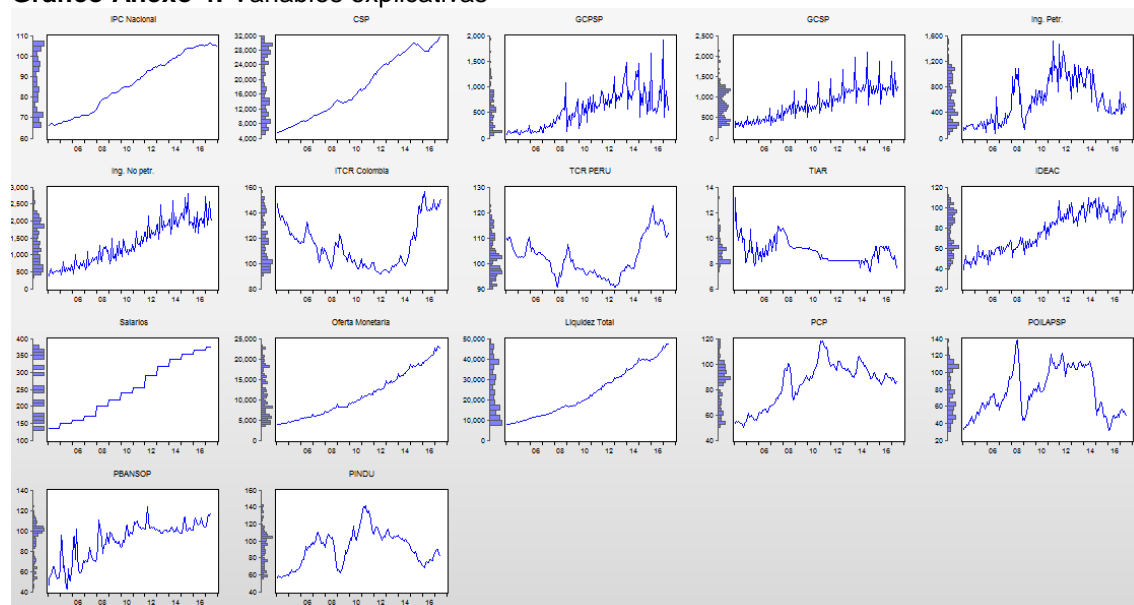
B.2. Variables explicativas

Tabla Anexo 2: Glosario de variables

Variable	Nomenclatura
Índice de precios al consumidor	IPC
Crédito al sector privado	CSP
Gasto de capital del sector publico	GCPSP
Gasto corriente del sector publico	GCSP
Ingreso petrolero	INP
Ingreso no petrolero	INNP
Índice de tipo de cambio real de Colombia	ITCR
Tipo de cambio real de Perú	TCR
Tasa de interés activa referencial	TIAR
Índice de desarrollo de actividad económica coyuntural	IDEAC
Salarios	Salarios
Oferta monetaria	M1
Liquidez Total	M2
Precio de productos primarios	PCP
Índice de precio de petróleo	POILAPSP
Índice del precio de banano	PBANSOP
Índice de precios de insumos industriales	PINDU

Elaboración Propia

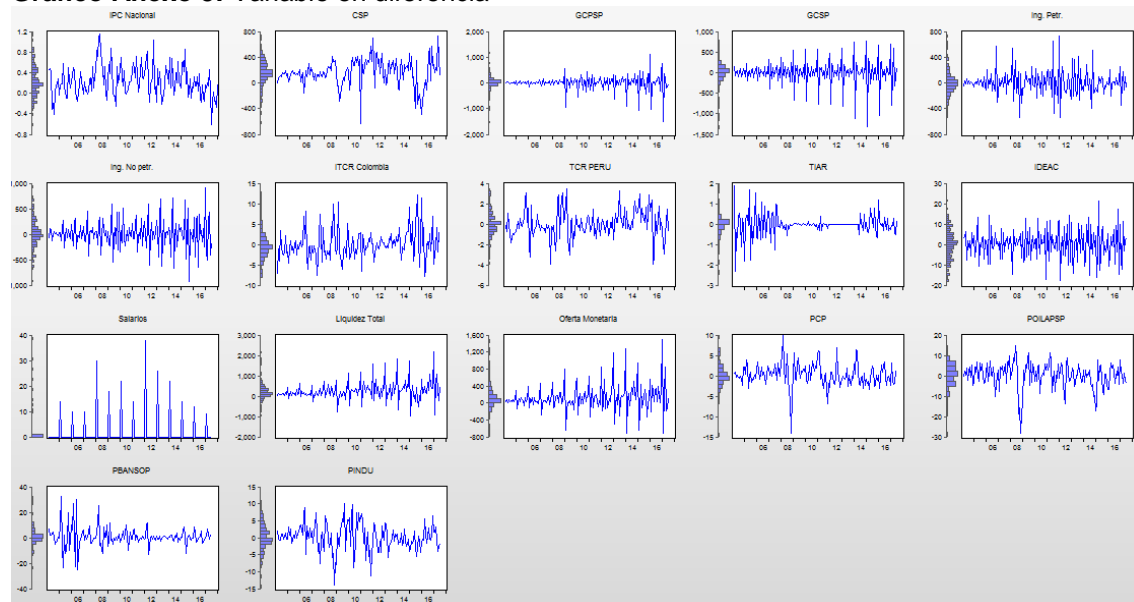
Gráfico Anexo 4: Variables explicativas



Fuente: Banco central, INEC, Banco Central de Colombia, Banco Central de Perú, FMI.
Elaboración propia.

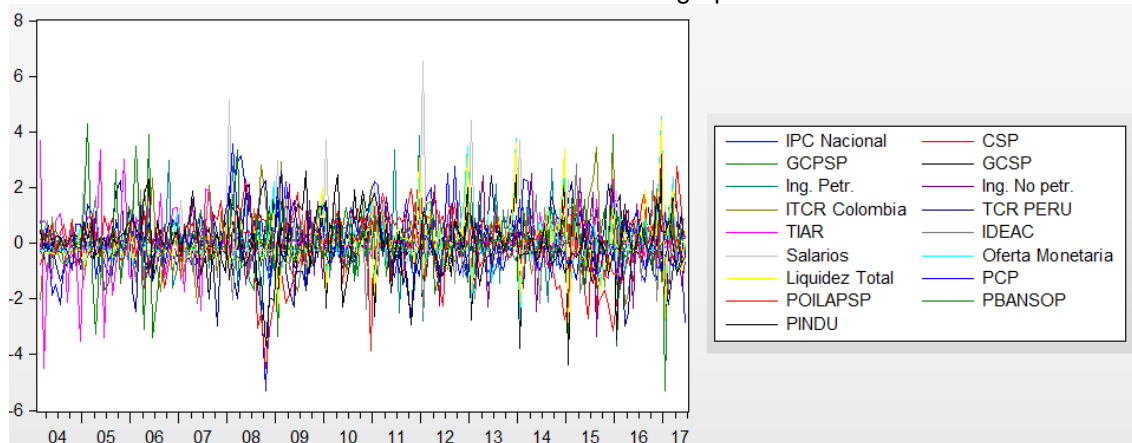
B.3. Tratamiento de Variables

Gráfico Anexo 5: Variable en diferencia



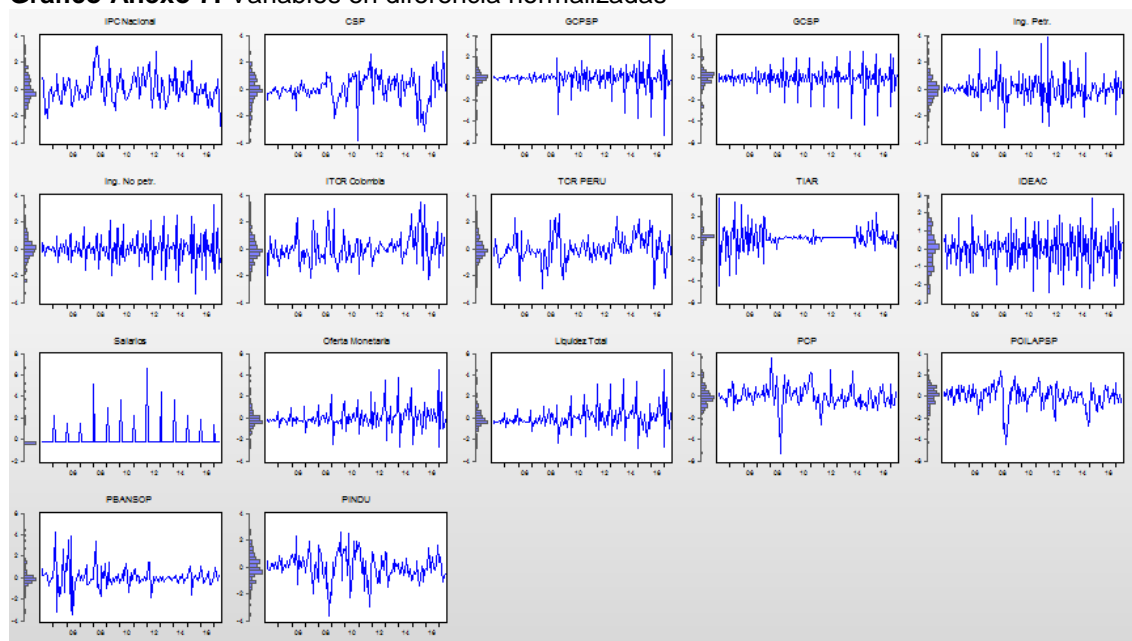
Fuente: Banco central, INEC, Banco Central de Colombia, Banco Central de Perú, FMI.
Elaboración propia.

Gráfico Anexo 6: Variables en diferencia normalizado agrupadas⁷⁷



Fuente: Banco central, INEC, Banco Central de Colombia, Banco Central de Perú, FMI.
Elaboración propia.

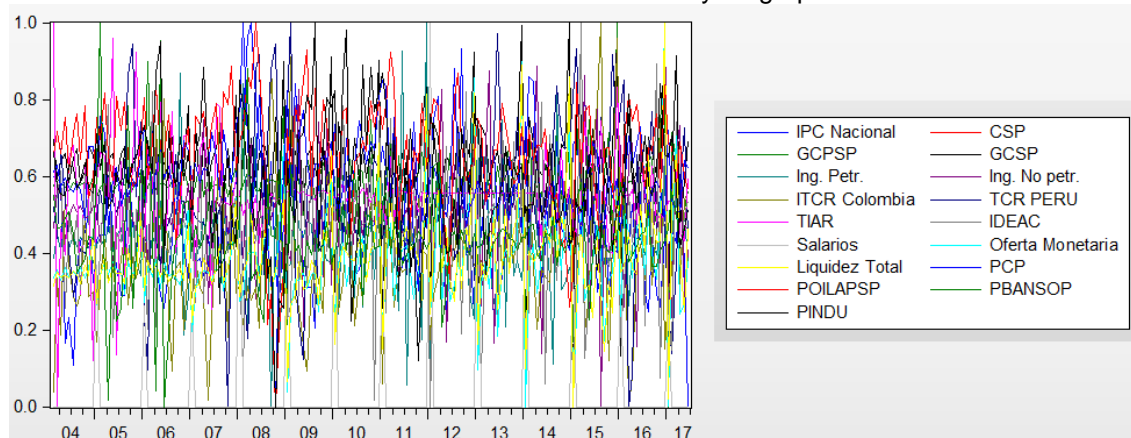
Gráfico Anexo 7: Variables en diferencia normalizadas



Fuente: Banco central, INEC, Banco Central de Colombia, Banco Central de Perú, FMI.
Elaboración propia.

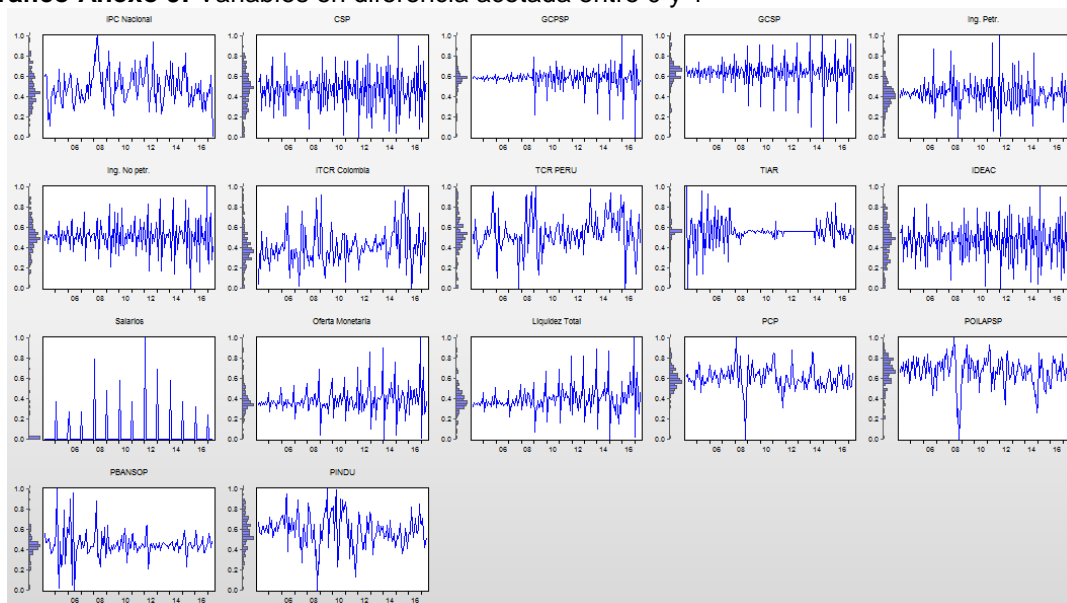
⁷⁷ Como se aprecia en la gráfica, las series ya no presentan problemas de estacionariedad ni de diferencia en escala y todas las variables siguen una distribución $n(0,1)$.

Gráfico Anexo 8: Variables en diferencia acotada entre 0 y 1 agrupadas



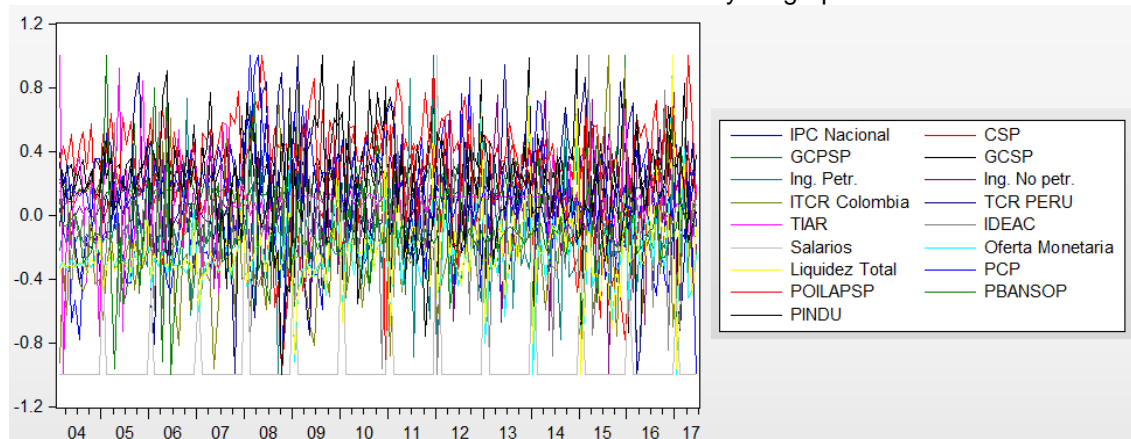
Fuente: Banco central, INEC, Banco Central de Colombia, Banco Central de Perú, FMI.
Elaboración propia.

Gráfico Anexo 9: Variables en diferencia acotada entre 0 y 1



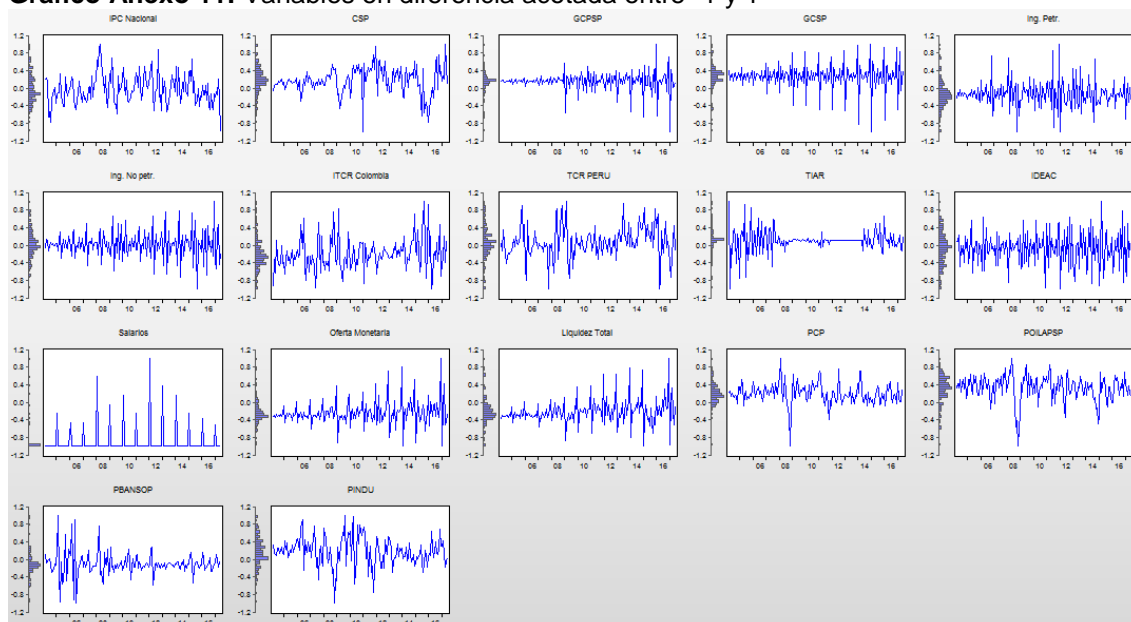
Fuente: Banco central, INEC, Banco Central de Colombia, Banco Central de Perú, FMI.
Elaboración propia.

Gráfico Anexo 10: Variables en diferencia acotada entre -1 y 1 agrupadas



Fuente: Banco central, INEC, Banco Central de Colombia, Banco Central de Perú, FMI.
Elaboración propia.

Gráfico Anexo 11: Variables en diferencia acotada entre -1 y 1



Fuente: Banco central, INEC, Banco Central de Colombia, Banco Central de Perú, FMI.
Elaboración propia.

B.4. Análisis de componentes principales

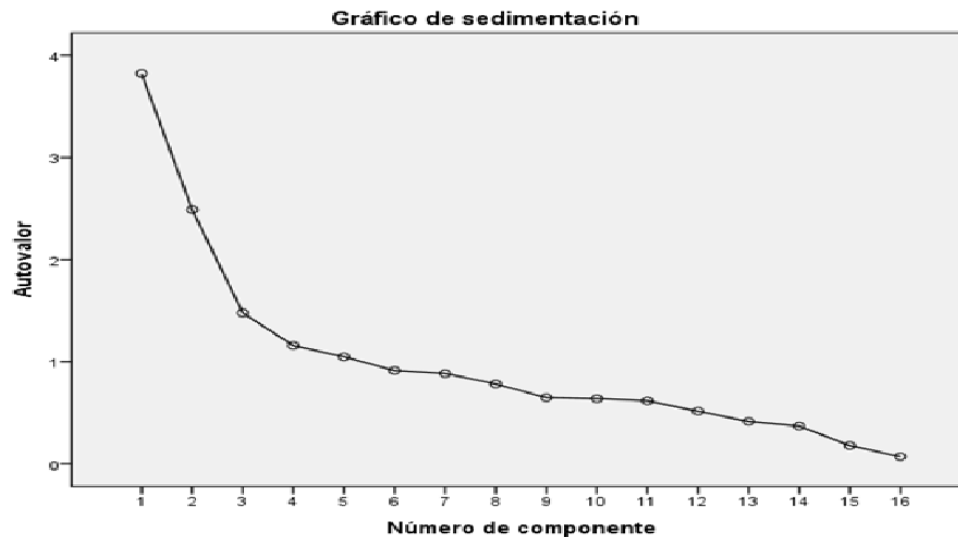
Gráfico Anexo 12: Prueba KMO y esfericidad de Bartlett

Medida Kaiser-Meyer-Olkin de adecuación de muestreo		,733
Prueba de esfericidad de Bartlett	Aprox. Chi-cuadrado	918,659
	gl	120
	Sig.	,000

Elaboración propia

Gráfico Anexo 13: Varianza explicada por cada componente y grafico de sedimentación⁷⁸.

Componente	Autovalores iniciales			Sumas de extracción de cargas al cuadrado			Sumas de rotación de cargas al cuadrado		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	3,823	23,891	23,891	3,823	23,891	23,891	3,716	23,227	23,227
2	2,488	15,548	39,439	2,488	15,548	39,439	2,504	15,647	38,874
3	1,477	9,230	48,670	1,477	9,230	48,670	1,317	8,234	47,109
4	1,156	7,223	55,892	1,156	7,223	55,892	1,273	7,956	55,065
5	1,048	6,547	62,439	1,048	6,547	62,439	1,180	7,374	62,439
6	,911	5,695	68,134						



Elaboración Propia.

Gráfico Anexo 14: Matriz de componentes

	Componente				
	1	2	3	4	5
ITCR Colombia	,274	-,512	-,322	-,308	,125
TCR PERU	-,022	-,610	-,136	-,306	,204
PCP	-,230	,701	-,225	-,076	-,054
POILAPSP	-,121	,773	,030	-,263	-,084
PBANSOP	-,042	,102	-,301	,649	,227
PINDU	-,193	,746	-,118	-,018	,062
Ing. Petr.	,087	,174	,328	-,315	,658
Ing. No petr.	,160	,152	,680	,021	,294
GCSP	,875	,157	-,235	-,067	-,008
GCPSP	,749	,137	-,146	-,089	-,087
Salarios	-,678	,001	,099	,082	,175
TIAR	,062	,001	-,376	,409	,523
IDEAC	,509	,090	,288	,049	,131
CSP	,204	-,225	,559	,412	-,276
Oferta Monetaria	,875	,132	-,049	,036	,013
Liquidez Total	,867	,157	,060	,104	,002

Método de extracción: análisis de componentes principales.

a. 5 componentes extraídos.

Elaboración Propia.

⁷⁸ Como se aprecia, el número de componentes a tomarse en cuenta son 5, mismos que representan el 62.44% de la varianza total de la muestra.

Gráfico Anexo 15: Matriz de componentes rotados⁷⁹

	Componente				
	1	2	3	4	5
ITCR Colombia	,225	-,562	,421	-,069	-,054
TCR PERU	-,111	-,620	,345	,049	-,093
PCP	-,069	,737	,222	-,076	,036
POILAPSP	,024	,779	,144	,119	-,216
PBANSOP	-,027	,099	-,114	-,148	,728
PINDU	-,053	,756	,150	,072	,101
Ing. Petr.	,022	,046	,201	,796	-,004
Ing. No petr.	,061	,057	-,388	,658	-,090
GCSP	,912	,000	,124	-,003	,059
GCPSP	,781	,012	,059	-,023	-,038
Salarios	-,691	,099	,005	,091	,106
TIAR	,050	-,061	,181	,083	,735
IDEAC	,457	-,031	-,227	,329	,002
CSP	,093	-,227	-,764	,015	-,075
Oferta Monetaria	,876	-,031	-,061	,094	,071
Liquidez Total	,857	-,006	-,183	,137	,072

Método de extracción: análisis de componentes principales.

Método de rotación: Varimax con normalización Kaiser. ^a

a. La rotación ha convergido en 6 iteraciones.

Elaboración Propia.

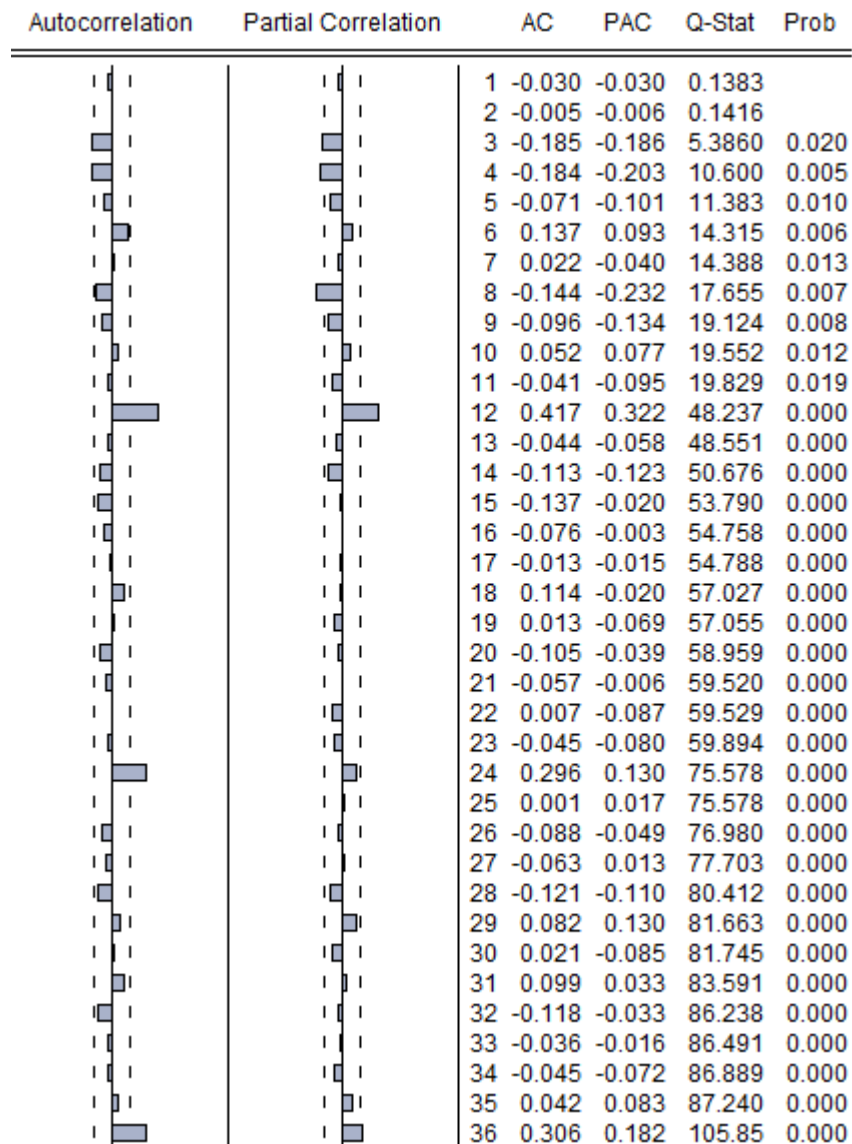
⁷⁹ La rotación varimax incrementa la interpretabilidad de los factores al reducir los coeficientes pequeños e incrementar los grandes, de manera tal que resulta más sencilla la interpretación de cada componente como un promedio ponderado de las variables (primer componente) y un promedio ponderado de *m* combinaciones de variables (*m* componentes restantes).

C. Resultados

C.1. Modelos ARIMA

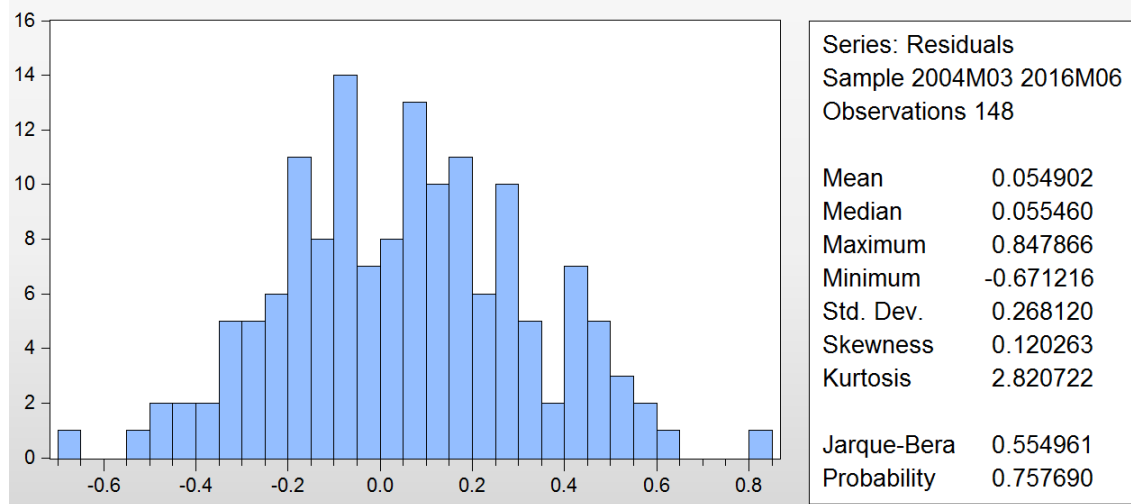
Modelo ARIMA (1, 1,1)

Gráfico Anexo 16: PAC y FAC Modelo ARIMA (1, 1,1)



Elaboración Propia

Gráfico Anexo 17: Histograma de errores modelo ARIMA (1, 1, 1)



Elaboración Propia.

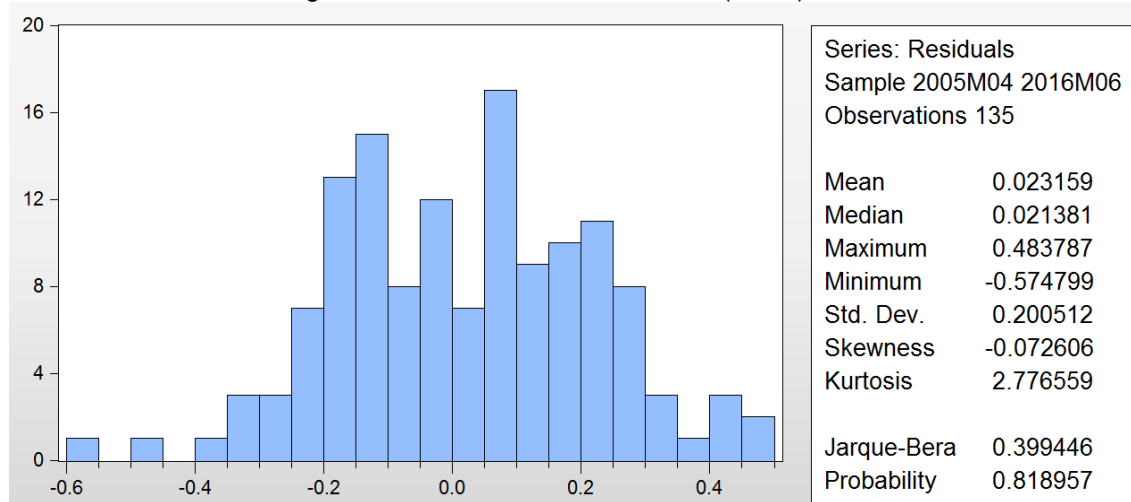
Modelo SARIMA (2, 1, 1)

Gráfico Anexo 18: Estimación modelo SARIMA (2, 1, 1)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(2)	0.426789	0.088075	4.845766	0.0000
MA(1)	0.475043	0.084965	5.591061	0.0000
SMA(12)	-0.905521	0.027682	-32.71180	0.0000
R-squared	0.645343	Mean dependent var		0.007667
Adjusted R-squared	0.639969	S.D. dependent var		0.338949
S.E. of regression	0.203378	Akaike info criterion		-0.325525
Sum squared resid	5.459886	Schwarz criterion		-0.260964
Log likelihood	24.97296	Hannan-Quinn criter.		-0.299289
Durbin-Watson stat	1.945903			
Inverted AR Roots	.65	-.65		
Inverted MA Roots	.99	.86-.50i	.86+.50i	.50-.86i
	.50+.86i	.00+.99i	-.00-.99i	-.48
	-.50+.86i	-.50-.86i	-.86+.50i	-.86-.50i
	-.99			

Elaboración Propia

Gráfico Anexo 19: Histograma de errores modelo SARIMA (2, 1, 1)



Elaboración Propia.

Gráfico Anexo 20: PAC y FAC Modelo SARIMA (2, 1, 1)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1	0.003	0.003	0.0011
		2	0.006	0.006	0.0063
		3	-0.032	-0.032	0.1474
		4	-0.065	-0.065	0.7481
		5	-0.116	-0.116	2.6743
		6	0.085	0.086	3.7093
		7	0.050	0.049	4.0733
		8	-0.089	-0.105	5.2324
		9	0.002	-0.009	5.2329
		10	0.174	0.186	9.7295
		11	-0.050	-0.033	10.100
		12	-0.040	-0.065	10.344
		13	-0.022	-0.038	10.416
		14	-0.159	-0.125	14.283
		15	-0.134	-0.101	17.038
		16	0.050	0.001	17.426
		17	-0.063	-0.104	18.056
		18	-0.066	-0.067	18.745
		19	0.075	0.047	19.653
		20	0.040	0.006	19.913
		21	-0.011	0.015	19.932
		22	0.082	0.067	21.023
		23	-0.134	-0.161	24.008
		24	-0.212	-0.164	31.481
		25	0.060	0.105	32.096
		26	-0.039	-0.108	32.350
		27	-0.052	-0.102	32.806
		28	-0.060	-0.129	33.437
		29	0.131	0.065	36.448
		30	-0.126	-0.114	39.227
		31	0.144	0.100	42.925
		32	0.048	-0.063	43.345
		33	-0.017	0.034	43.394
		34	-0.015	0.107	43.436
		35	0.132	0.074	46.678
		36	-0.007	0.002	46.688

Elaboración Propia

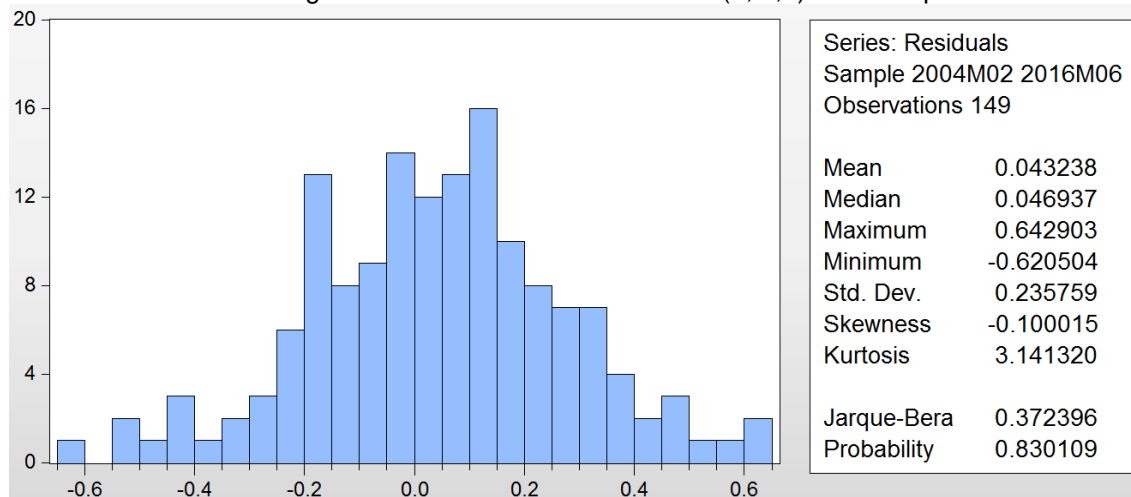
Modelo SARIMA (1, 1, 1) Mediano Plazo

Gráfico Anexo 21: Estimación modelo SARIMA (1, 1, 1) mediano plazo

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.757335	0.084088	9.006455	0.0000
SAR(12)	0.490398	0.070113	6.994355	0.0000
MA(1)	-0.241000	0.133983	-1.798734	0.0741
SIGMASQ	0.057079	0.006579	8.676390	0.0000
R-squared	0.296016	Mean dependent var		0.264664
Adjusted R-squared	0.281451	S.D. dependent var		0.285706
S.E. of regression	0.242185	Akaike info criterion		0.054054
Sum squared resid	8.504753	Schwarz criterion		0.134697
Log likelihood	-0.027054	Hannan-Quinn criter.		0.086818
Durbin-Watson stat	2.014320			
Inverted AR Roots	.94	.82+.47i	.82-.47i	.76
	.47+.82i	.47-.82i	.00-.94i	-.00+.94i
	-.47-.82i	-.47+.82i	-.82+.47i	-.82-.47i
	-.94			
Inverted MA Roots	.24			

































































Elaboración Propia.

Gráfico Anexo 22: Histograma de errores modelo SARIMA (1, 1, 1) mediano plazo



Elaboración Propia.

Gráfico Anexo 23: PAC y FAC Modelo SARIMA (1, 1,1) mediano plazo

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.047	-0.047	0.3383	
		2	0.045	0.043	0.6542	
		3	-0.120	-0.116	2.8707	
		4	-0.108	-0.123	4.6915	0.030
		5	-0.095	-0.099	6.0916	0.048
		6	0.085	0.072	7.2260	0.065
		7	0.064	0.055	7.8703	0.096
		8	-0.044	-0.083	8.1804	0.147
		9	0.006	-0.012	8.1871	0.225
		10	0.067	0.102	8.9158	0.259
		11	-0.024	0.002	9.0084	0.342
		12	-0.130	-0.163	11.788	0.226
		13	-0.042	-0.067	12.085	0.279
		14	-0.104	-0.071	13.876	0.240
		15	-0.066	-0.090	14.596	0.264
		16	0.054	-0.021	15.091	0.302
		17	-0.049	-0.119	15.499	0.345
		18	0.035	0.008	15.705	0.402
		19	-0.031	-0.030	15.872	0.462
		20	-0.016	-0.065	15.919	0.530
		21	-0.009	-0.011	15.933	0.597
		22	0.011	0.011	15.956	0.660
		23	-0.085	-0.109	17.259	0.636
		24	0.026	-0.014	17.381	0.688
		25	0.049	0.040	17.809	0.717
		26	-0.033	-0.090	18.007	0.757
		27	-0.017	-0.088	18.058	0.800
		28	-0.091	-0.150	19.611	0.767
		29	0.117	0.095	22.160	0.680
		30	-0.082	-0.086	23.428	0.662
		31	0.157	0.031	28.135	0.457
		32	-0.022	-0.049	28.228	0.506
		33	-0.021	-0.018	28.316	0.554
		34	-0.056	-0.054	28.929	0.573
		35	0.066	0.007	29.778	0.579
		36	0.136	0.161	33.479	0.444

Elaboración Propia.

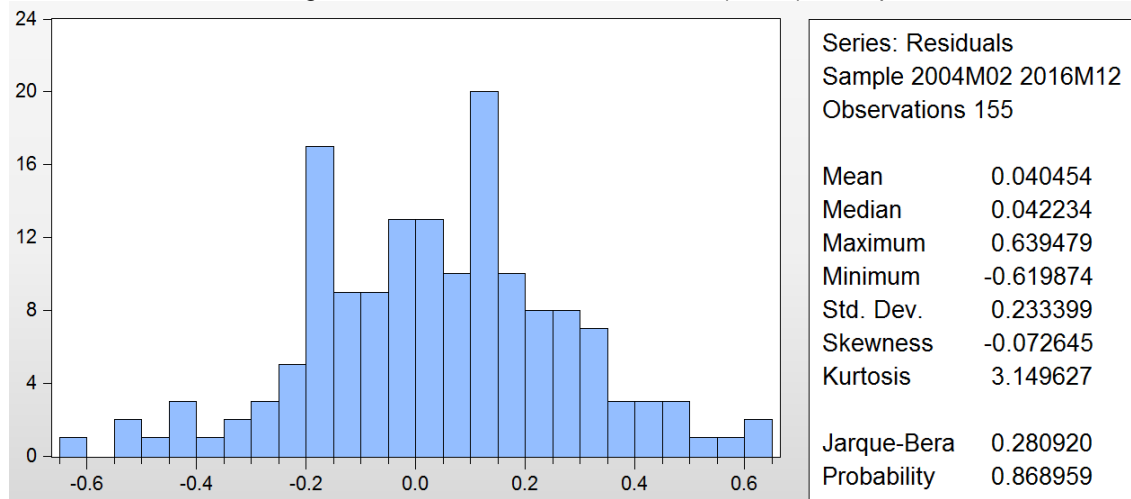
Modelo SARIMA (1, 1,1) Corto Plazo

Gráfico Anexo 24: Estimación modelo SARIMA (1, 1, 1) corto plazo

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.756577	0.082813	9.135944	0.0000
SAR(12)	0.498239	0.068400	7.284148	0.0000
MA(1)	-0.248897	0.131428	-1.893787	0.0602
SIGMASQ	0.055760	0.006250	8.921735	0.0000
R-squared	0.319205	Mean dependent var		0.253297
Adjusted R-squared	0.305679	S.D. dependent var		0.287118
S.E. of regression	0.239243	Akaike info criterion		0.028318
Sum squared resid	8.642849	Schwarz criterion		0.106858
Log likelihood	1.805363	Hannan-Quinn criter.		0.060219
Durbin-Watson stat	2.011474			
Inverted AR Roots	.94	.82+.47i	.82-.47i	.76
	.47+.82i	.47-.82i	.00-.94i	-.00+.94i
	-.47+.82i	-.47-.82i	-.82-.47i	-.82+.47i
	-.94			
Inverted MA Roots	.25			

Elaboración Propia.

Gráfico Anexo 25: Histograma de errores modelo SARIMA (1, 1,1) corto plazo



Elaboración Propia.

Gráfico Anexo 26: PAC y FAC Modelo SARIMA (1, 1,1) corto plazo

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.042	-0.042	0.2843	
		2 0.039	0.037	0.5231	
		3 -0.104	-0.101	2.2630	
		4 -0.106	-0.117	4.0576	0.044
		5 -0.097	-0.102	5.5917	0.061
		6 0.094	0.083	7.0262	0.071
		7 0.064	0.059	7.6969	0.103
		8 -0.043	-0.079	8.0018	0.156
		9 0.016	-0.000	8.0437	0.235
		10 0.068	0.102	8.8088	0.267
		11 -0.026	0.003	8.9203	0.349
		12 -0.124	-0.152	11.530	0.241
		13 -0.024	-0.046	11.631	0.310
		14 -0.110	-0.075	13.716	0.249
		15 -0.062	-0.083	14.375	0.277
		16 0.057	-0.008	14.942	0.311
		17 -0.058	-0.117	15.539	0.342
		18 0.021	0.003	15.617	0.408
		19 -0.049	-0.055	16.050	0.449
		20 0.006	-0.031	16.057	0.520
		21 -0.020	-0.008	16.129	0.584
		22 0.016	0.002	16.175	0.646
		23 -0.081	-0.098	17.375	0.628
		24 0.026	0.005	17.499	0.680
		25 0.042	0.050	17.832	0.716
		26 -0.036	-0.091	18.082	0.753
		27 -0.004	-0.062	18.086	0.799
		28 -0.105	-0.141	20.185	0.737
		29 0.116	0.104	22.772	0.646
		30 -0.067	-0.069	23.653	0.649
		31 0.142	0.029	27.612	0.485
		32 -0.029	-0.047	27.775	0.530
		33 -0.015	-0.010	27.819	0.580
		34 -0.074	-0.067	28.917	0.574
		35 0.070	0.027	29.924	0.572
		36 0.126	0.152	33.167	0.459

Elaboración Propia.

C.2. Modelos VAR

Modelo VAR exploratorio

Tabla Anexo 3: Especificación VAR exploratorio

Variables consideradas	Todas
Tratamiento variables	Diferenciación
Numero de rezagos	1
Máximo de rezagos	8
Residuos ruido blanco	No

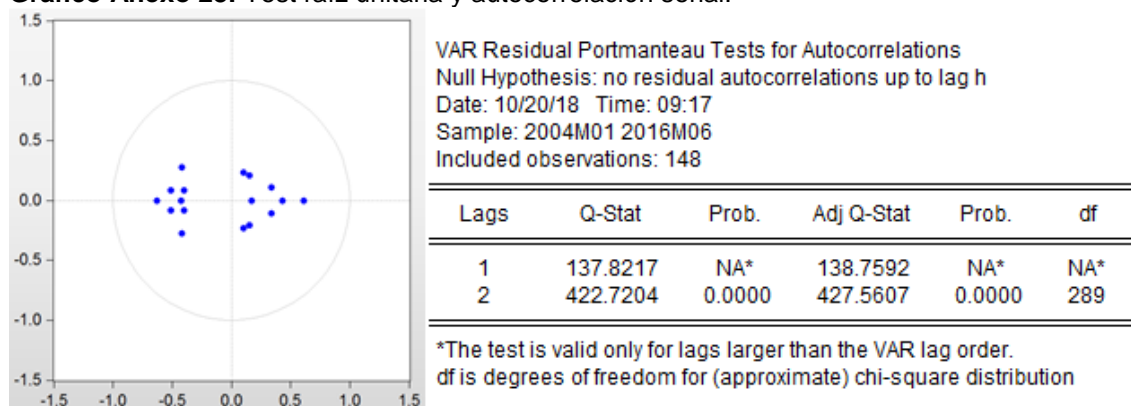
Elaboración Propia.

Gráfico Anexo 27: Resultados VAR exploratorio

R-squared	0.427023
Adj. R-squared	0.352095
Sum sq. resids	6.902262
S.E. equation	0.230422
F-statistic	5.699127
Log likelihood	16.83396
Akaike AIC	0.015757
Schwarz SC	0.380283
Mean dependent	0.263412
S.D. dependent	0.286265
Determinant resid covariance (dof adj.)	4.59E+36
Determinant resid covariance	5.06E+35
Log likelihood	-9653.697
Akaike information criterion	134.5905
Schwarz criterion	140.7874

Elaboración Propia.

Gráfico Anexo 28: Test raíz unitaria y autocorrelación serial.



Elaboración Propia.

Modelo VAR variables de la revisión de la literatura

Tabla Anexo 4: Especificación VAR variables de la revisión de la literatura 1 rezago

Variables consideradas	TCR, ITCR, TIAR, PCP, Salarios
Tratamiento variables	Diferenciación
Numero de rezagos	1
Máximo de rezagos	>14
Residuos ruido blanco	No
RECM	3.14

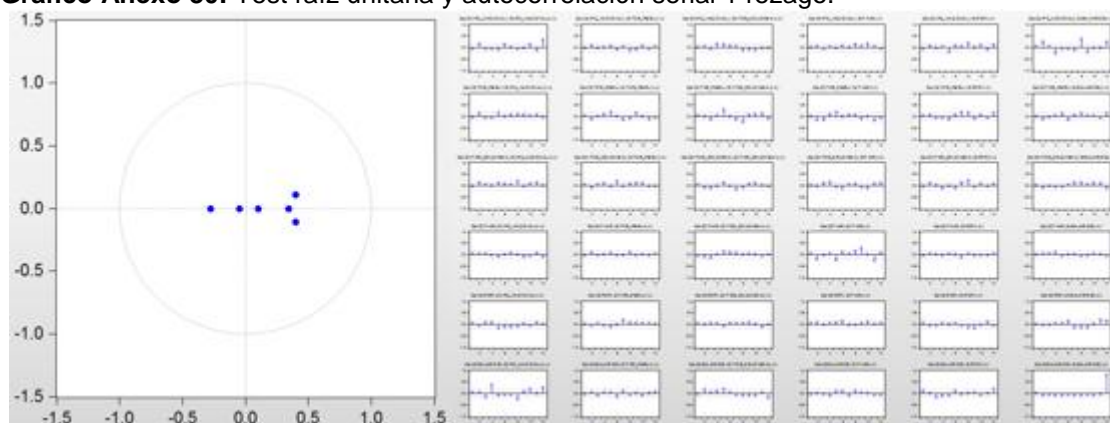
Elaboración Propia.

Gráfico Anexo 29: Resultados VAR variables de la revisión de la literatura 1 rezago.

R-squared	0.229353
Adj. R-squared	0.196560
Sum sq. resids	9.283450
S.E. equation	0.256593
F-statistic	6.993876
Log likelihood	-5.098455
Akaike AIC	0.163493
Schwarz SC	0.305253
Mean dependent	0.263412
S.D. dependent	0.286265
Determinant resid covariance (dof adj.)	38.12196
Determinant resid covariance	28.50491
Log likelihood	-1507.923
Akaike information criterion	20.94491
Schwarz criterion	21.79547

Elaboración Propia.

Gráfico Anexo 30: Test raíz unitaria y autocorrelación serial 1 rezago.



Elaboración Propia.

Gráfico Anexo 31: Especificación VAR variables de la revisión de la literatura 14 rezagos.

Variables consideradas	TCR, ITCR, TIAR, PCP, Salarios
Tratamiento variables	Diferenciación
Numero de rezagos	14
Máximo de rezagos	>14
Residuos ruido blanco	Si
RECM	3.90

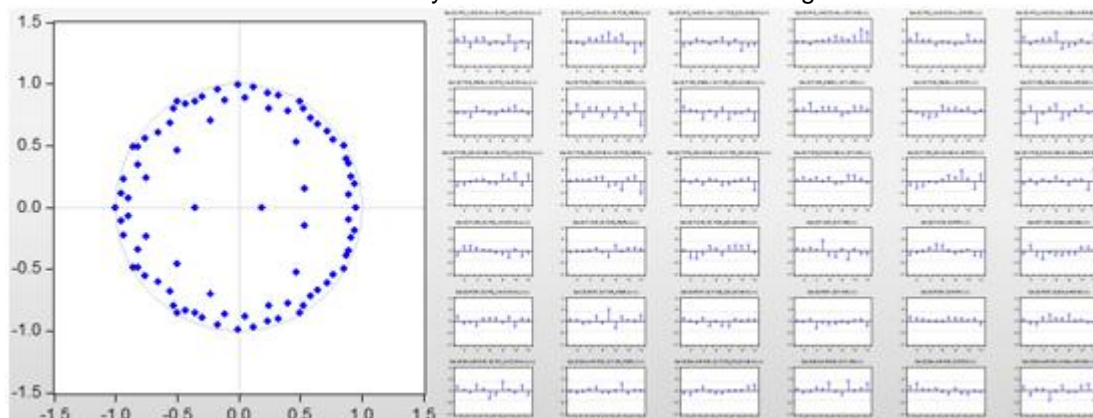
Elaboración Propia.

Gráfico Anexo 32: Resultados VAR variables de la revisión de la literatura 14 rezagos.

R-squared	0.818406
Adj. R-squared	0.513328
Sum sq. resids	1.946424
S.E. equation	0.197303
F-statistic	2.682615
Log likelihood	94.59476
Akaike AIC	-0.142145
Schwarz SC	1.687102
Mean dependent	0.280778
S.D. dependent	0.282824
Determinant resid covariance (dof adj.)	3.513753
Determinant resid covariance	0.009070
Log likelihood	-831.8995
Akaike information criterion	19.87999
Schwarz criterion	30.85547

Elaboración Propia.

Gráfico Anexo 33: Test raíz unitaria y autocorrelación serial 14 rezagos.



Elaboración Propia.

Modelo VAR componentes principales

Tabla Anexo 5: Especificación VAR componentes principales.

Variables consideradas	Componentes principales
Tratamiento variables	Análisis CP
Numero de rezagos	1
Máximo de rezagos	>14
Residuos ruido blanco	Si
RECM	2.18

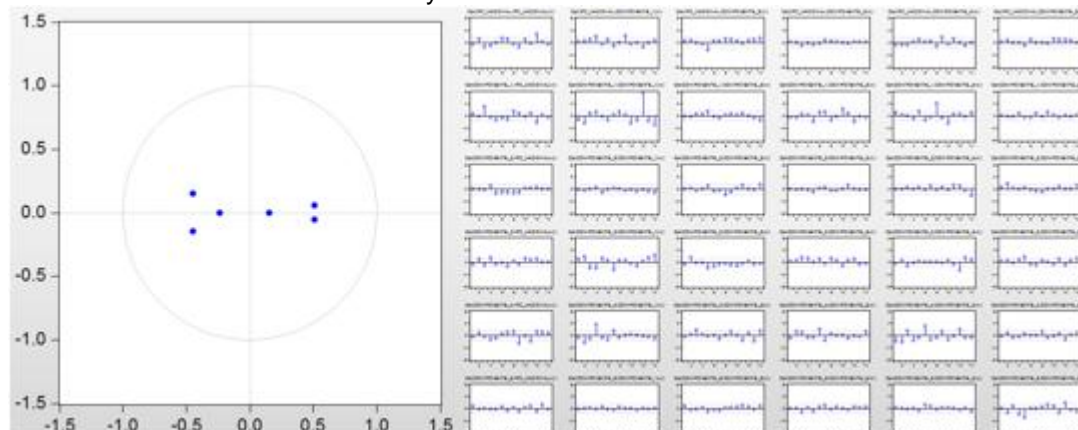
Elaboración Propia.

Gráfico Anexo 34: Resultados VAR componentes principales.

R-squared	0.338350
Adj. R-squared	0.319842
Sum sq. resids	92.16193
S.E. equation	0.802801
F-statistic	18.28157
Log likelihood	-174.9517
Akaike AIC	2.431780
Schwarz SC	2.533037
Mean dependent	0.100584
S.D. dependent	0.973426
Determinant resid covariance (dof adj.)	0.239960
Determinant resid covariance	0.209140
Log likelihood	-724.2200
Akaike information criterion	10.05703
Schwarz criterion	10.46206

Elaboración Propia.

Gráfico Anexo 35: Test raíz unitaria y autocorrelación serial



Elaboración Propia.

C.3. Modelos RNA

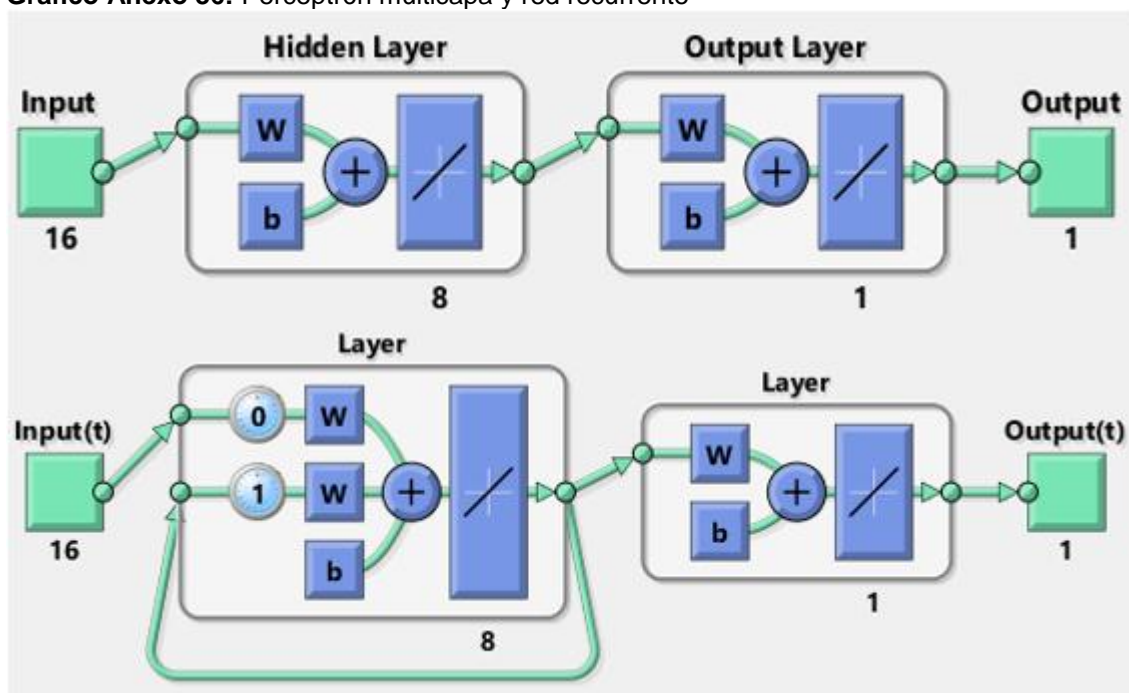
Tabla Anexo 6: Descripción modelos RNA.

Topología modelo RNA	Nomenclatura
Datos en diferencia normalizados, red recurrente, 1 capa oculta, función lineal ambas capas, 32 neuronas	RNA 1
Datos en diferencia normalizados, red recurrente, 1 capa oculta, función lineal ambas capas, 8 neuronas	RNA 2
Datos en diferencia normalizados, red recurrente, 1 capa oculta, función lineal ambas capas, 4 neuronas	RNA 3
Datos en diferencia normalizados, red recurrente, 2 capa oculta, función lineal en tres capas, 8 neuronas capa uno, 4 neuronas capa dos.	RNA 4
Datos en diferencia normalizados, red recurrente, 2 capa oculta, función lineal en tres capas, 8 neuronas capa uno, 8 neuronas capa dos.	RNA 5
Datos en diferencia normalizados, red recurrente, 2 capa oculta, función lineal en tres capas, 16 neuronas capa uno, 4 neuronas capa dos.	RNA 6
Datos en diferencia normalizados, red perceptrón, 1 capa oculta, función lineal ambas capas, 8 neuronas.	RNA 7
Datos en diferencia normalizado acotado 0 y 1, red recurrente, 1 capa oculta, función sigmoidea capa oculta, lineal capa de salida, 8 neuronas.	RNA 8
Datos en diferencia normalizado acotado 0 y 1, red recurrente, 1 capa oculta, función sigmoidea dos capas, 8 neuronas.	RNA 9
Datos en diferencia normalizado acotado -1 y 1, red recurrente, 1 capa oculta, función tangente hiperbólica capa oculta, lineal capa de salida, 8 neuronas.	RNA 10
Datos en diferencia normalizado acotado -1 y 1, red recurrente, 1 capa oculta, función tangente hiperbólica ambas capas.	RNA 11
Datos componentes principales, red recurrente, 1 capa oculta, función lineal ambas capas, 8 neuronas	RNA 12
Datos completos en diferencia normalizados, red recurrente, 1 capa oculta, función lineal ambas capas, 8 neuronas	RNA 13

Elaboración Propia.

Modelo RNA determinación de arquitectura.

Gráfico Anexo 36: Perceptrón multicapa y red recurrente



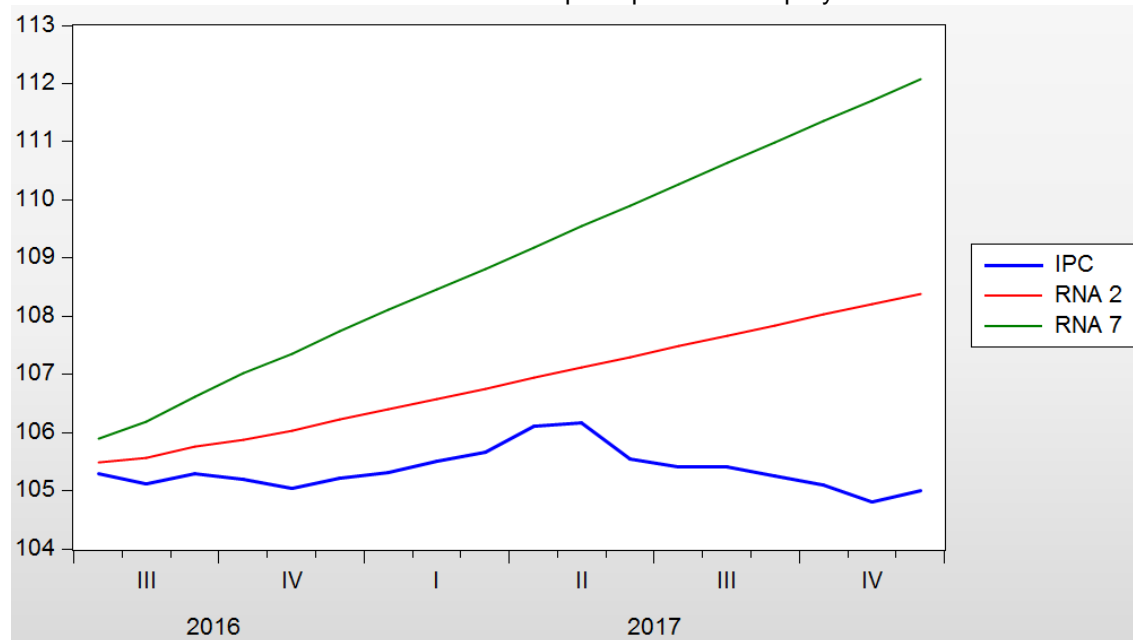
Elaboración Propia.

Gráfico Anexo 37: Pronósticos modelos RNA perceptrón multicapa y red recurrente.

Modelo	ECM	RECM	MPE	MAPE	σ_{MAPE}	U Theil
RNA 2	3.259163	1.805315	1.4341%	1.4341%	0.0105%	0.008506
RNA 7	16.96006	4.118259	3.4523%	3.4523%	0.0401%	0.019210

Elaboración Propia.

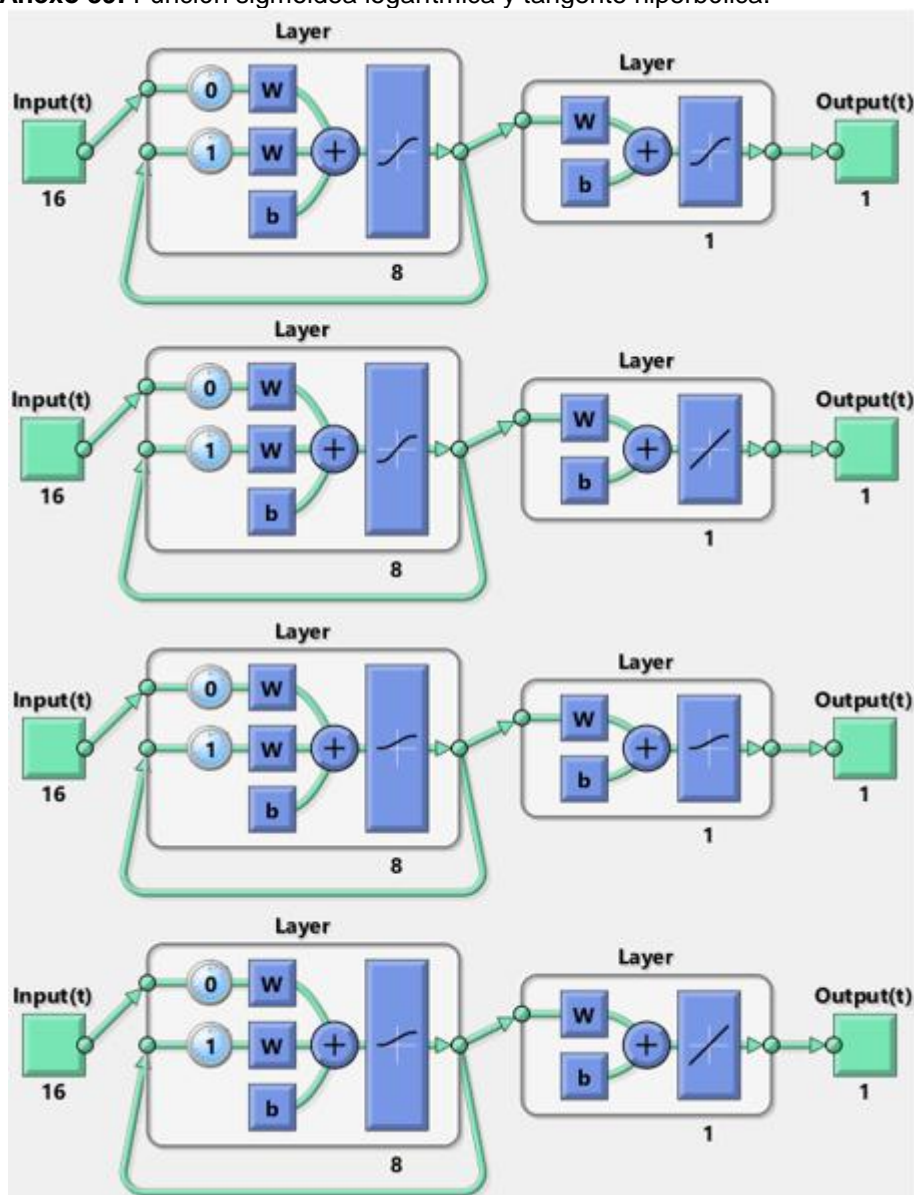
Gráfico Anexo 38: Pronósticos modelos RNA perceptrón multicapa y red recurrente.



Elaboración Propia.

Modelo RNA determinación de la función de activación.

Gráfico Anexo 39: Función sigmoidea logarítmica y tangente hiperbólica.



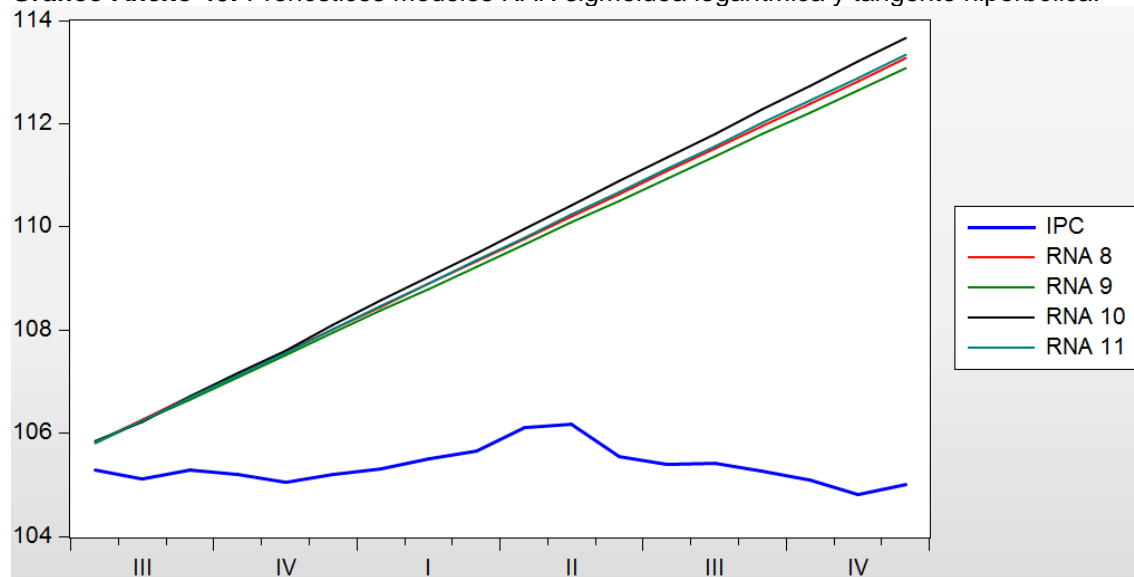
Elaboración Propia.

Tabla Anexo 7: Resultados modelos RNA sigmoidea logarítmica y tangente hiperbólica.

Modelo	ECM	RECM	MPE	MAPE	σ_{MAPE}	U Theil
RNA 8	22.88538	4.783866	3.9777%	3.9777%	0.0581%	0.022257
RNA 9	21.78962	4.667935	3.8780%	3.8780%	0.0555%	0.021729
RNA 10	25.04751	5.004749	4.1464%	4.1464%	0.0650%	0.023265
RNA 11	23.24084	4.818850	4.0007%	4.0007%	0.0595%	0.022417

Elaboración Propia.

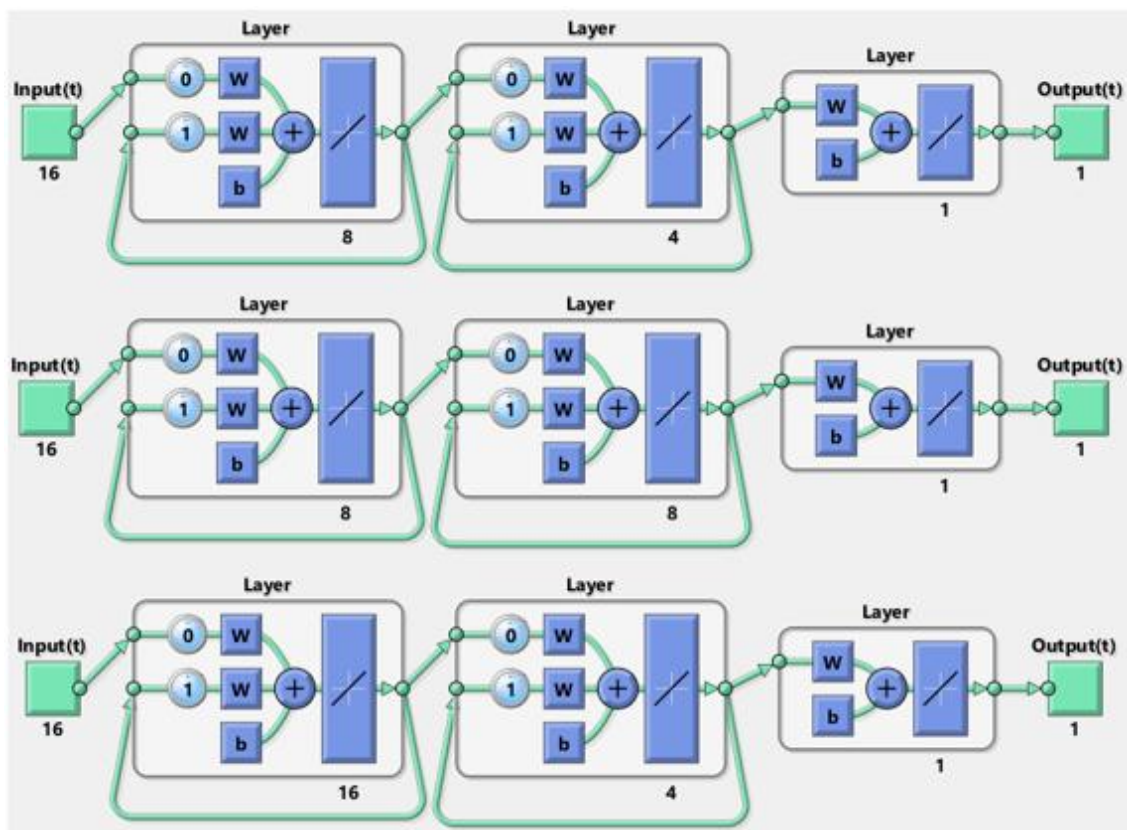
Gráfico Anexo 40: Pronósticos modelos RAN sigmoidea logarítmica y tangente hiperbólica.



Elaboración Propia.

Modelo RNA determinación de las capas ocultas.

Gráfico Anexo 41: Arquitectura con 2 capas ocultas.



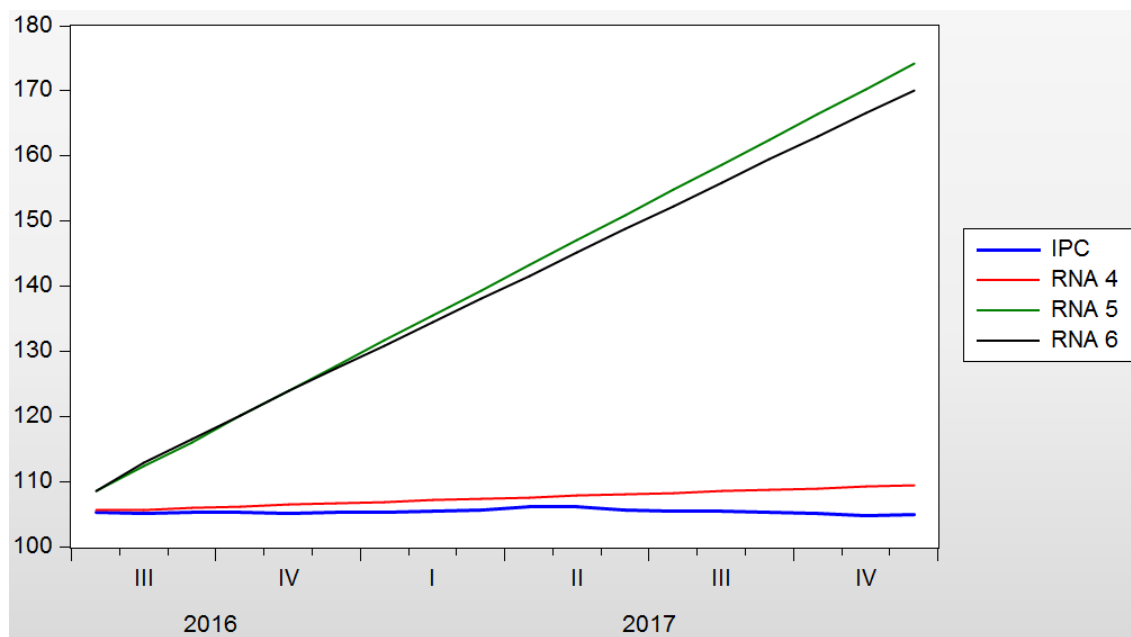
Elaboración Propia.

Tabla Anexo 8: Resultados modelos RNA dos capas ocultas.

Modelo	ECM	RECM	MPE	MAPE	σ_{MAPE}	U Theil
RNA 4	6.029236	2.455450	2.0067%	2.0067%	0.0165%	0.011536
RNA 5	1694.463	41.16385	34.1245%	34.1245%	4.4587%	0.165929
RNA 6	1530.493	39.12152	32.6600%	32.6600%	3.8347%	0.158800

Elaboración Propia.

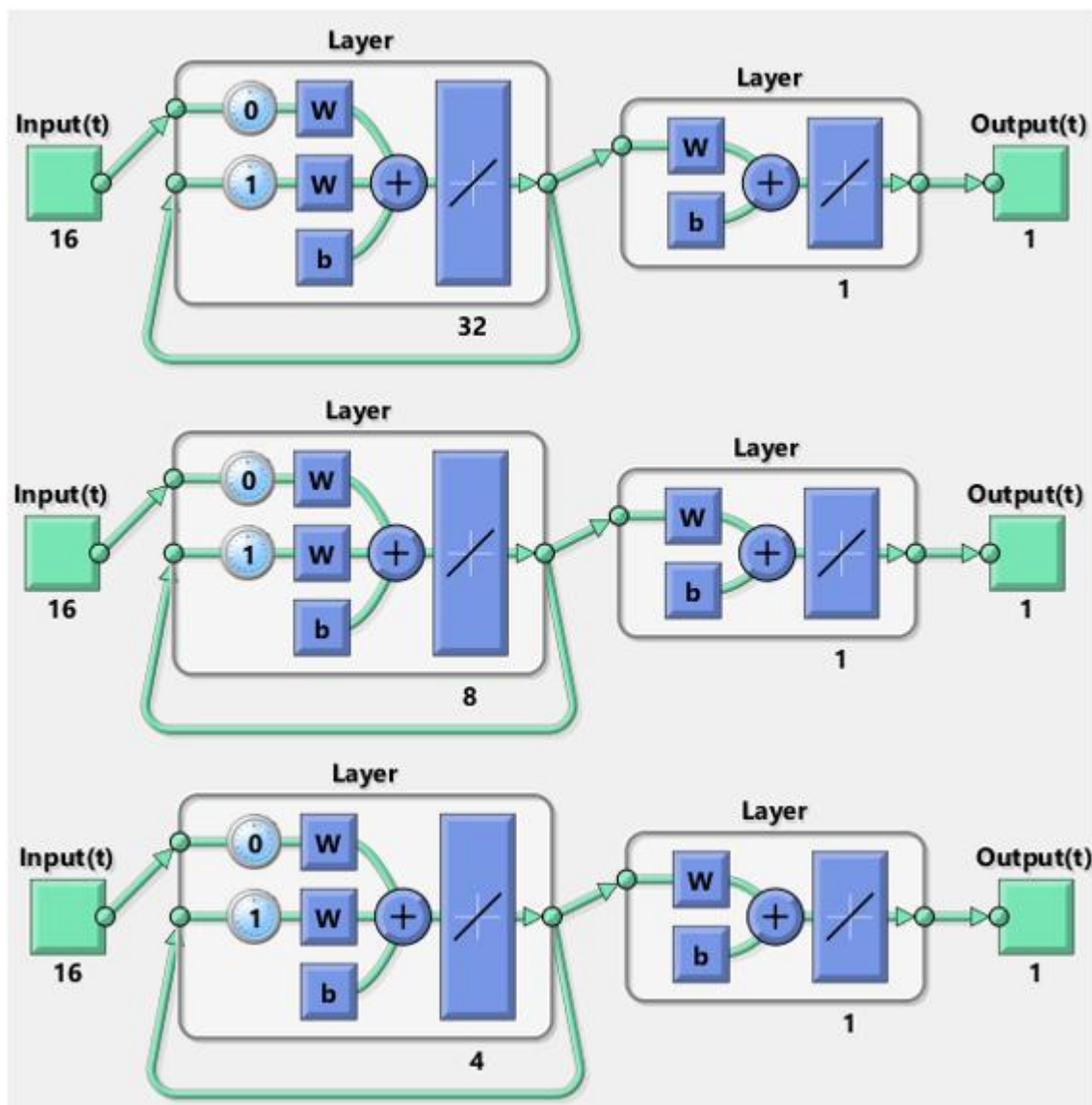
Gráfico Anexo 42: Pronósticos modelos RNA dos capas ocultas.



Elaboración Propia.

Modelo RNA determinación del número de neuronas de cada capa.

Gráfico Anexo 43: Reducción del número de neuronas ocultas.



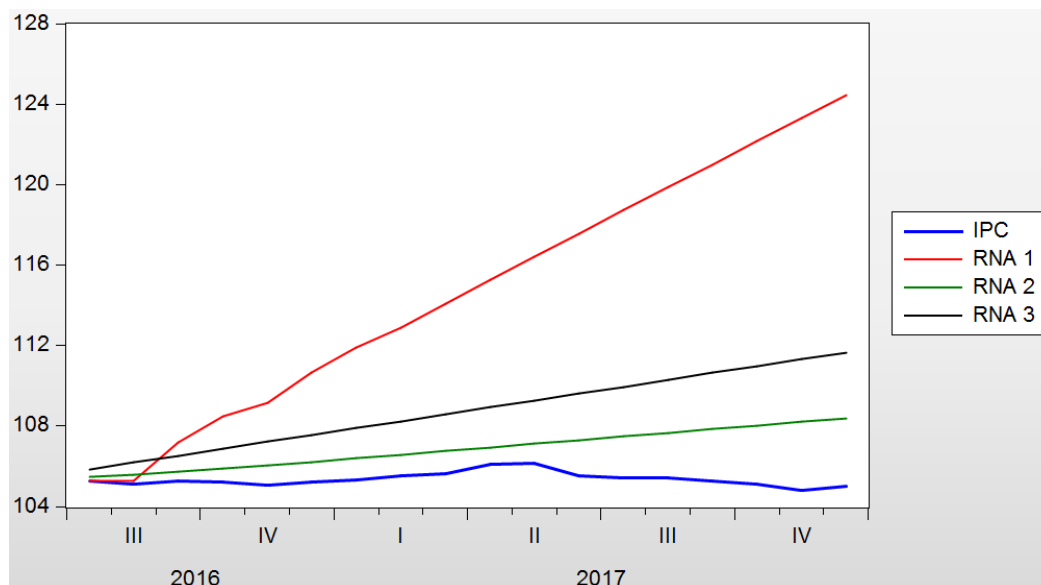
Elaboración Propia.

Gráfico Anexo 44: Resultados modelos RNA diferente número de neuronas.

Modelo	ECM	RECM	MPE	MAPE	σ_{MAPE}	U Theil
RNA 1	122.5773	11.07146	8.8201%	8.8212%	0.39345%	0.050286
RNA 2	3.259163	1.805315	1.4341%	1.4341%	0.0105%	0.008506
RNA 3	14.86585	3.855626	3.2237%	3.2237%	0.0360%	0.018006

Elaboración Propia.

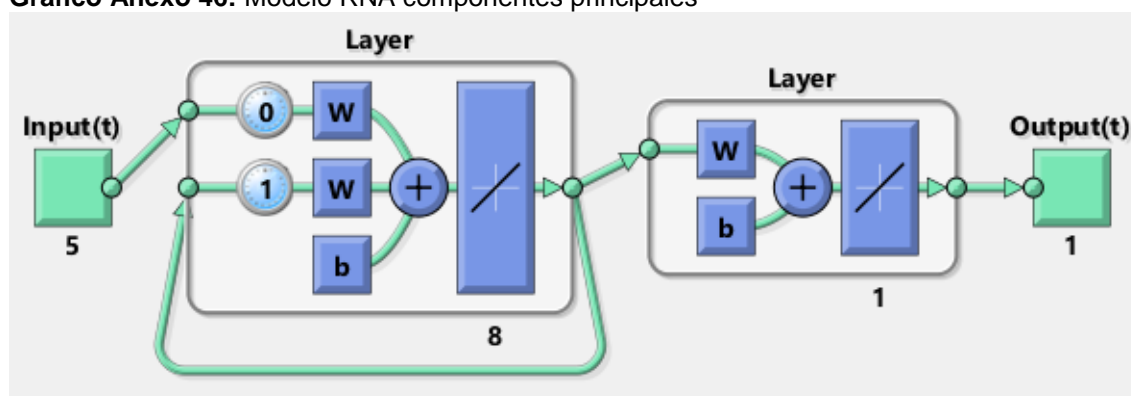
Gráfico Anexo 45: Pronostico modelos RNA diferente número de neuronas.



Elaboración Propia.

Modelo RNA componentes principales.

Gráfico Anexo 46: Modelo RNA componentes principales



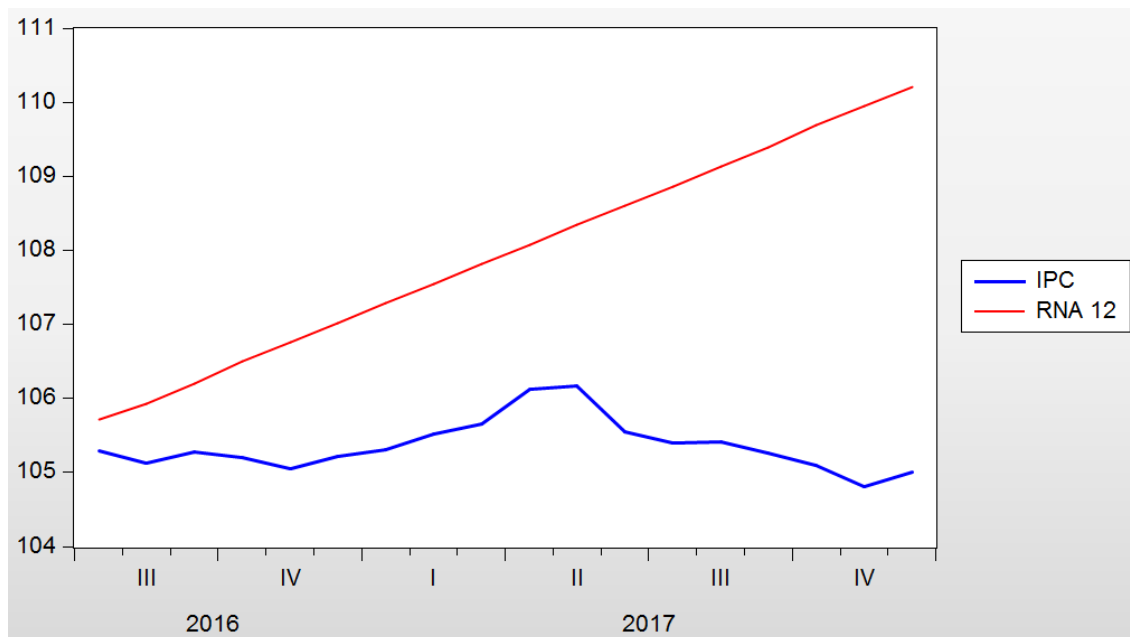
Elaboración Propia.

Tabla Anexo 9: Resultados modelo RNA componentes principales.

Modelo	ECM	RECM	MPE	MAPE	σ_{MAPE}	U Theil
RNA 12	7.377010	2.716065	2.2225%	2.2225%	0.0176%	0.012747

Elaboración Propia.

Gráfico Anexo 47: Pronostico modelo RNA componentes principales.



Elaboración Propia.

D. Protocolo

1. Resumen de la Propuesta.

Una de las variables macroeconómicas de gran relevancia para las autoridades económicas de un país es el Índice de Precios al Consumidor (IPC), misma que nos permite llevar a cabo el análisis sobre el impacto que tiene la inflación sobre los niveles de consumo de los hogares, además de obtener una forma de medir la calidad de vida de la población a través de la capacidad de consumo de las familias y la capacidad adquisitiva que tienen los salarios (Instituto Nacional de Estadísticas y Censos, 2015), de la misma manera esta variable resulta fundamental al momento de llevar a cabo el ajuste anual de las rentas, las pensiones o los salarios, aspecto por el cual el pronóstico preciso de esta variable resulta de gran importancia para la fijación de metas de política económica por parte de las autoridades.

Dada la capacidad que tiene esta variable para reflejar el desempeño económico de una nación, el pronóstico acertado de la misma resulta fundamental al momento de tomar decisiones en lo que respecta a la política económica debido a que el control de la inflación dentro de un rango saludable (inferior a las dos cifras), es un prerequisite indispensable para que una economía consiga el desarrollo sostenido dentro de una economía de libre mercado (Erraez J. P., 2005). Otro aspecto por el cual resulta fundamental un pronóstico acertado del IPC es que este le permite a las autoridades económicas y a los sectores empresariales nacionales el anticipar en cierta medida las modificaciones en cuanto a los salarios y de esta manera enfocar de mejor forma sus políticas salariales por un lado y sus niveles de inversión por el otro, con la finalidad de mejorar sus rendimientos y optimizar sus recursos.

Para llevar a cabo el pronóstico de variables de series de tiempo, como es el IPC y las variables explicativas de mismo, los modelos utilizados de forma regular son los

Modelos Auto regresivos Integrados de Media Móvil (ARIMA), desarrollados en el año de 1970 por George Box y Gwilym Jenkins (George Box, 1976) y los Modelos de Vectores Auto regresivos (VAR), desarrollados en el año de 1980 por Christopher Sims (Sims, 1980). Sin embargo el desarrollo de nuevas herramientas informáticas han permitido la creación de modelos que dadas ciertas circunstancias, podrían arrojar resultados mucho más precisos que los modelos tradicionales, como lo son los modelos de Redes Neuronales Artificiales (RNA), mismos que se utilizarán en este estudio y cuyos resultados serán contrastados con los modelos tradicionales ARIMA y VAR, con la finalidad de poder determinar si el uso de las RNA se justifica como una alternativa de los modelos tradicionales o brinda mejores resultados como un complemento de los mismos.

2. Identificación y Justificación.

Tal es la importancia del IPC para todas las economías a nivel mundial que organizaciones internacionales como la OIT, CEPAL, OCDE, FMI y BM⁸⁰ llevan a cabo un constante monitoreo de este indicador macroeconómico y las correspondientes recomendaciones respecto al mismo, todo con la finalidad de asegurar el correcto desenvolvimiento del sistema económico de los diferentes países a nivel mundial.

Dichos organismos definen al IPC como: *“El IPC es un índice que mide el ritmo al que los precios de los bienes y servicios de consumo cambian de un mes (o de un trimestre) a otro”* (Organización Internacional de Trabajo; Fondo Monetario Internacional; Organización para la Cooperación y el Desarrollo Económico; Oficina Estadística de las Comunidades Europeas; Organización de las Naciones Unidas; Banco Mundial, 2006). Dado que este índice se compone de los precios recogidos en los comercios de diversos productos de consumo y se obtiene un promedio porcentual del cambio en el precio de los mismos de un periodo a otro, este nos ayuda a medir la inflación o cambio en los precios de los bienes de consumo y como estos evolucionan entre periodos y a través de ello nos permite realizar un sondeo de la estabilidad económica, poder adquisitivo y el desarrollo económico de una nación.

Dado que uno de los factores principales para el establecimiento de los precios dentro del mercado es la demanda de los bienes, el incremento saludable y sostenido de la inflación (entendiéndose como saludable a cifras de inflación inferiores al 10%), nos muestra un estado dinámico de la economía, donde esta se encuentra demandando una cantidad mayor de bienes que se refleja en el incremento de su valor monetario (SEMPLADES, 2008). Por otro lado cuando en una economía se presentan procesos deflacionarios o una inflación muy cercana a cero, se entiende este como un síntoma de desaceleración en la actividad económica, es decir que tanto las familias como las empresas han reducido sus niveles de consumo, este panorama resulta muy preocupante para las autoridades económicas ya que en el largo plazo puede resultar en problemas de desempleo, debilitamiento del sistema productivo nacional, pérdida de competitividad en los mercados internacionales y la falta de inversión extranjera por considerar esta una economía poco dinámica.

⁸⁰ OIT (Organización Mundial del Trabajo), CEPAL (La Comisión Económica Para América Latina y El Caribe), OCDE (La Organización para la Cooperación y el Desarrollo Económico), FMI (El Fondo Monetario Internacional), BM (Banco Mundial)

Son estas las razones por las que la predicción del IPC es una herramienta fundamental para las autoridades nacionales en la creación de políticas económicas que se encarguen de controlar esta variable en caso de ser necesario y poder tomar las medidas necesarias en cuanto a las diferentes aristas capaces de modificar la misma. De tal manera que se pueda hasta cierto punto anticipar los efectos que tendría el afectar cualquiera de estas variables sobre la evolución del IPC, todo con miras a fomentar el desarrollo económico sostenido y con ello mejorar las condiciones de vida de la ciudadanía.

El acceso a nuevas herramientas computacionales por parte de la población, las mismas que ofrecen una gran capacidad de procesamiento de información, se ha empezado a potencializar la aplicación de nuevas herramientas como la de Redes Neuronales Artificiales para la detección de patrones en el análisis de datos y el desarrollo de predicciones en base de los resultados obtenidos, los mismos que son muy prometedores en cuanto a sus propiedades y capacidades y que los convierten en una alternativa atractiva dentro de los campos donde han sido utilizados.

3. Revisión de la Literatura.

El primer trabajo que se centró en los modelos para series de tiempo fue el trabajo realizado por George Box y Gwilym Jenkins en el año de 1970, mismo que es detallado en su libro titulado *"Times Series Analysis: Forecasting and Control"*, en este trabajo se lleva a cabo un procedimiento que utiliza los modelos ARMA y ARIMA basado en tres etapas para realizar una evaluación comparativa de modelos de predicción con datos de serie de tiempo. Las tres etapas que utilizan Box y Jenkins empiezan con la selección del modelo, donde se prueba la estacionariedad, estacionalidad y auto correlación (total o parcial), la segunda etapa consiste en la estimación de los parámetros a través de métodos estadísticos que mejor se ajusten al modelo y por último se lleva a cabo la comprobación del modelo a través del ensayo (George Box, 1976), este procedimiento se empezaría a conocer como la metodología de Box-Jenkins.

Otro trabajo llevado a cabo para el pronóstico de la inflación es el realizado por Andrew Ang, Geert Bekaert y Min Wei en el cual utilizan una serie de modelos ARIMA y modelos lineales para encontrar cuál de estos es el que ofrece mejores resultado en el pronóstico fuera de muestra de la inflación (Andrew Ang, 2006). En este trabajo se obtiene que los modelos ARIMA utilizados son los que mejores resultados han dado de todos los modelos, esto debido a que la inflación presenta procesos estocásticos, los modelos contra los cuales se lleva a cabo el contraste de resultados son los modelos de caminata aleatoria anual, regresiones basadas en la curva de Phillips y los modelos MCO (Andrew Ang, 2006).

Otra de las modelaciones más comunes utilizadas en variables de serie de tiempo con fines de pronóstico son los modelos VAR, los mismos que fueron ampliamente tratados por Christopher A. Sims en su trabajo titulado *"Macroeconomics and Reality"* en el año de 1980, donde se recurre a un modelo VAR para el estudio del ciclo de los negocios en un periodo de 10 años (Sims, 1980). Dicha propuesta de modelación para variables de serie de tiempo es motivado por el gran número de limitaciones a manera de supuestos que presentan las modelaciones de series de tiempo usadas tradicionalmente hasta ese entonces, tales como el hecho de no tomar en consideración la presencia de variables explicativas que pueden influir en la serie modelada, sino que solamente se consideraba

la serie en sí misma y el comportamiento que esta presentase. Por lo tanto tampoco se tomaba en consideración la fuerte presencia de correlación entre estas variables ni mucho menos la existencia de componentes cíclicos que podrían afectar la modelación de las series de interés, dichos supuestos impedían que estos modelos se adapten de manera correcta a la realidad. Precisamente el modelo VAR le permite dejar de lado algunos supuestos que resultaban muy restrictivos, brindando mejores resultados en el análisis del ciclo de negocios.

Con respecto a los modelos de RNA varias son las disciplinas que aplican esta herramienta para encontrar patrones dentro de los datos y de esta manera poder realizar buenos pronósticos, sin embargo, la RNA que tiene mayor relevancia para el pronóstico de las variables económicas es el denominado perceptron multicapa, cuyos fundamentos para llevar a cabo su implementación son tratados de manera bastante clara por Jose Castro, en su trabajo de titulación denominado *“Fundamentos para la implementación de red neuronal perceptron multicapa mediante software”* y que brinda los lineamientos necesarios para llevar a cabo la correcta implementación de este modelo (Garcia, 2006)

Otra aplicación de los modelos de RNA es la presentada por Bruce Baker y Craig Richards, en el cual utilizan un modelo de RNA y realizan un contraste de estos resultados con modelos de regresión lineal convencionales, dichas metodologías son aplicadas para el pronóstico del gasto en educación (Bruce Baker, 1999). En este estudio los resultados obtenidos por las RNA no difieren mucho de los resultados obtenidos con los modelos más convencionales. En el análisis posterior llevado a cabo se concluye que esto se da debido a que dos de las variables que se usan como variables explicativas dentro del modelo, son altamente lineales, razón por la cual las bondades de los modelos RNA (la mayor adaptabilidad en variables no lineales) no es aprovechada, por lo que los resultados obtenidos no difieren mucho entre ambos modelos.

Dentro del campo económico el uso de las RNA han sido explotadas principalmente dentro de los mercados financieros (tipo de interés y de cambio, indicadores bursátiles, evolución del precio de acciones). Los usos de estos modelos han sido motivados principalmente porque permite sortear varias restricciones de los modelos tradicionales como el ARIMA, restricciones tales como los problemas de multicolinealidad en series multi periodo o el hecho de tener que llevar a cabo la especificación previa de la relación funcional que se piense existe entre variables (Alberto Aragón, 2014). Otro de los campos dentro de la economía donde se ha llevado a cabo la utilización de RNA como herramienta de pronóstico alternativo a los modelos estadísticos clásicos es en la creación de modelos para la predicción de quiebres empresariales, obteniendo mejores resultados en los pronósticos que los modelos Logit y Probit utilizados.

4. Contexto del Problema y Preguntas de Investigación.

Como ya se ha mencionado anteriormente en los apartados desarrollados a lo largo de esta propuesta metodológica, la presente investigación busca obtener un pronóstico efectivo del IPC del Ecuador utilizando como base los datos obtenidos en el periodo post dolarización dadas las implicancias que tiene este importante indicador macroeconómico, tanto para las autoridades económicas nacionales como para el sector privado empresarial, los mismos que toman en consideración este indicador para llevar

a cabo importantes ajustes en decisiones de política económica por un lado y las proyecciones de financiamiento, costos, ingresos e inversiones por el otro.

El desarrollo de la ciencia y la tecnología durante la última década ha sido exponencial, sobre todo en los aspectos informáticos, los mismos que le han permitido a la población común tener una gran capacidad de procesamiento de datos y con ello el resurgimiento de técnicas que necesitan una gran cantidad de recursos (motivo por el cual su desarrollo de los modelos de RNA ha sido muy limitado). Tal es el caso de las RNA las mismas que pese a que no son una herramienta relativamente nueva (en comparación con otras herramientas utilizadas tradicionalmente) su inmersión dentro del campo científico ha sido bastante discreto y ha sido solo durante la última década que la mayoría de ciencias han empezado a incorporar estas metodologías de forma exhaustiva para llevar a cabo sus análisis y pronósticos, obteniendo resultados muy prometedores.

La disciplina económica no puede rezagarse en cuanto a los avances tecnológicos y la aparición de nuevas herramientas para llevar a cabo pronósticos de series de tiempo, mismos que son imprescindibles dentro de la ciencia económica, muchos investigadores e importantes institutos han llevado a cabo estudios utilizando RNA⁸¹. La presente propuesta metodológica lo que busca es llevar a cabo el pronóstico del IPC para la economía ecuatoriana a través de RNA y contrastar los resultados obtenidos con modelos que son utilizados tradicionalmente para este propósito, como lo son los modelos ARIMA y RNA,

4.1 Preguntas de investigación.

1. ¿Resulta válida la aplicación de modelos de RNA (perceptron multicapa) para el pronóstico del IPC en el Ecuador post dolarización?
2. ¿Cuál de los modelos a utilizarse (RNA, ARIMA, VAR) ofrecen una mejor capacidad predictiva tanto en el corto como en el mediano plazo, utilizando la/las técnicas de evaluación pertinentes?
3. ¿La utilización de las RNA se justifica de mejor manera como una herramienta sustituta o complementaria a los modelos tradicionales ARIMA y VAR?

5. Marco Teórico.

5.1 Marco de Referencia

Uno de los primeros economistas en abordar la inflación de una manera formal fue Irvin Fisher a través de su teoría cuantitativa del dinero donde plantea la fórmula $M \times V \equiv P \times Y$ donde M representa la cantidad de dinero en la economía, V la velocidad de circulación del mismo, P la inflación e Y es el Producto Interno Bruto real (PIBr) (Gregorio, 2007), en dicha formulación Fisher nos da a entender que la inflación corresponde a un fenómeno monetario y que el correcto manejo de la política monetaria resulta el mecanismo más eficaz para controlar la inflación. De la misma manera John Maynard Keynes aborda la inflación desde otro punto de vista, centrándose principalmente en la oferta y demanda, concluyendo que existen tres fuerzas

⁸¹ Esta herramienta ha sido utilizada en varios estudios llevados a cabo por los bancos centrales de varios países, como el caso de la utilización de RNA para la construcción de mapas de pobreza para Costa Rica, la utilización de los modelos RNA para llevar a cabo la predicción de la inflación para Colombia, entre otros.

determinantes de la misma, primero asegura que el incremento en la demanda agregada hace que la oferta se vea desbordada, lo que ocasiona en primera instancia un fuerte incremento en los precios. Un segundo factor se relaciona al incremento de los costos de producción los mismos que el productor decide traspasar al consumidor a través de un incremento en el precio de los bienes finales. Por ultimo otro factor es un proceso de inflación estructural, en el cual se produce una especie de espiral inflacionaria de la cual es muy difícil salir (Keynes, 1924). Es por esto que Keynes era un defensor de la intervención del estado como un mecanismo eficaz de control de la inflación, a través del control de la oferta y la demanda de bienes. Por último, una de las teorías más utilizadas para explicar la inflación es la llevada a cabo por Milton Friedman, donde se indica que la inflación responde a las expectativas de los agentes económicos respecto a su ingresos tanto actuales como futuros, diversos activos presentes en la economía y sus tipos de interés y los niveles de ahorro y consumo presentes en una economía, los mismo que a la larga determinan la cantidad de dinero en circulación y la velocidad con la que este se mueve, es decir que la inflación es un fenómeno netamente monetario (Argandoña, 1990).

Pese a que en el largo plazo la inflación siempre es un fenómeno monetario, en el corto plazo los factores monetarios resultan ser menos determinantes que otros factores (Pablo Garcia, 2003), lo que hasta cierto punto resulta favorable para la realización del estudio en el Ecuador, país que ha perdido parte de su política monetaria desde el año 2000. En el año 2000 Grubisic y Manteiga llevan a cabo un trabajo para el Banco Central de Argentina donde buscan realizar la predicción de la inflación en base a un grupo de agregados monetarios, pese a que para estos años dicha economía había perdido gran parte de sus herramientas monetarias en pos de la estabilidad inflacionaria (Elena Grubisic, 2000). En dicho modelo se toman en cuenta como variables exógenas al PIB, EMI, spread del EMBI, tasa de depósitos a plazo fijo, los índices en bolsa, la inflación rezagada un periodo, además de variables dummies que recojan ciertos efectos estacionales. Los resultados obtenidos fueron que la variable más importante para explicar la inflación es la misma inflación rezagada, las demás variables resultaron apenas significativas dentro de la modelación llevada a cabo por Grubisic y Manteiga.

Varios son los estudios realizados con respecto a la dinámica de la inflación que se ha vivido en el país a partir del año 2000, tanto por el Banco Central así como por investigadores de los departamentos de economía de las diferentes universidades a nivel nacional. Todos estos estudios rescatan un factor importante dentro de la economía nacional que es la dolarización y el quiebre estructural que ocurrió en los años posteriores a la misma. Previo a la dolarización era común tener escenarios de inflación desmedida, lo que limitaba la capacidad predictiva de los modelos estadísticos utilizados, sin embargo tras la dolarización en el año 2000 y con ello la perdida de la posibilidad de realizar emisiones de dinero, la inflación se estabilizo considerablemente, por lo que los modelo a aplicarse empezaron a ofrecer resultado mucho más precisos.

En cuanto a trabajos relacionados con la economía ecuatoriana, Juan Pablo Erraez lleva a cabo en el año 2005 una investigación denominada “*El proceso inflacionario en el Ecuador: un análisis de sus determinantes con modelos ARIMA y Vectores Auto regresivos*” (Erraez J. P., 2005). Otro estudio disponible en el repositorio del BC es

aquel realizado por Gachet, Maldonado y Pérez en el año 2008 denominado *“Determinantes de la inflación en una economía dolarizada: El caso ecuatoriano”* (Ivan Gachet, 2008).

Dentro de la investigación realizada por Erraez utiliza dos modelos para analizar los factores determinantes de la inflación en el Ecuador en el periodo 2000-2004, tomando una especial consideración de la adopción del dólar como moneda nacional en el año 2000 y la posterior etapa de ajustes de los principales agregados macroeconómicos ocurridos tras este periodo. Las variables explicativas utilizadas por Erraez como determinantes de la inflación son: la tasa de interés activa referencial, egresos del presupuesto del gobierno central, índice de actividad económica y crédito al sector privado (Erraez J. P., 2005), para realizar el análisis los modelos utilizados son un modelo VAR y un modelo ARIMA. Los resultados obtenidos por Erraez son que la variable que mejor explica la inflación, es la misma inflación rezagada un periodo, mientras que las variables relacionadas con el crédito al sector privado y el índice de actividad económica resultan ser apenas significativas dentro del modelo.

En cuanto a la investigación de Gachet, Maldonado y Pérez, deciden optar por un modelo de vectores auto regresivos estructurales (SVAR) como herramienta para determinar los factores que inciden de manera directa como determinantes de la inflación durante el periodo 2000-2007. Como variables explicativas de la inflación se utiliza una medida de los precios internacionales, tipo de cambio, variables de política pública, el clima comercial, costo de fletes y transporte y eventos políticos que pudieron haber afectado la inflación (Ivan Gachet, 2008). En este estudio se logra determinar que las variables que mayor influencia han tenido para determinar la inflación del Ecuador son las variables relacionadas a los precios internacionales, el tipo de cambio de los socios comerciales y las políticas públicas implementadas.

Por otro lado, pese a que la aplicación de RNA dentro de la disciplina económica ha sido bastante limitada, se han realizado importantes estudios utilizando esta herramienta para el pronóstico de series de tiempo, con resultados bastante prometedores. Tal es el caso del análisis realizado por (Nakamura, 2004) donde utiliza un modelo de redes neuronales para llevar a cabo el pronóstico del IPC de los Estados Unidos de Norteamérica, para validar la capacidad predictiva de este modelo se utilizan dos metodologías de predicción (detección anticipada y técnicas de pre procesamiento) las mismas que se contrastan con los resultados obtenidos con un modelo autoregresivo univariado. Los resultados obtenidos en el estudio realizado por Nakamura son que los modelos RNA resultaron tener mayor eficiencia en las predicciones tanto para corto como para mediano plazo, mientras que en el largo plazo demostraron una menor eficiencia.

Una de las instituciones con más alto renombre internacional como es el Banco Central Europeo (BCE) también se ha visto motivado a la utilización de los modelos de RNA para el pronóstico de series de tiempo. Uno de los estudios más importantes llevados a cabo por el BCE se realizó en el año 2004 donde se buscó medir la eficacia de la predicción de la inflación basada en la curva de Phillips a través de la utilización de RNA (Paul McNelis, 2004). En dicho documento se discute la importancia de la

aplicación de modelos de RNA en series de datos que mantienen una relación lineal. Los resultados más destacables a los que llega McNelis es el buen desempeño que tienen los modelos RNA frente a modelos utilizados tradicionalmente y destacan que estos muestran un mejor desempeño cuando dentro del periodo de análisis existen cambios estructurales o cierto grado de incertidumbre.

Otra importante propuesta sobre el uso de RNA ya no como una alternativa a los modelos tradicionales, sino como un complemento para estos es el estudio denominado *“Neural Networks for Macroeconomics Forecasting: A Complementary Approach to Linear Regression Models”* (Gonzalez, 2007). Dentro de este estudio Gonzalez señala que la validez de la aplicación de modelos RNA no se basa en que estos substituyan a los modelos clásicos, sino que estos funcionan como un complemento de los mismo, potencializando de esta manera la capacidad de pronóstico de las series tratadas más que si los dos modelos trabajaran por separado. Dentro de las principales conclusiones a las que se llegan en este estudio es el hecho de encontrar cuatro posibles debilidades que pueden evitar que los modelos de RNA tengan una buena capacidad predictiva las mismas que se pueden resumir en la manera en la que se manejen los datos y en la utilización previa de un correcto modelo de regresión⁸².

5.2 Objetivos e Hipótesis

5.2.1 Objetivo General

Realizar un pronóstico del Índice de Precios al Consumidor para la economía ecuatoriana periodo 2000-2017 a través de la utilización de modelos ARIMA, VAR y RNA.

5.2.2 Objetivos Específicos

- Llevar a cabo la evaluación de la capacidad de ajuste en la modelación del IPC ofrecida por cada uno de los modelos propuestos, utilizando para ello las herramientas de evaluación pertinentes.
- Identificar cuál de los modelos propuestos (RNA, ARIMA, VAR) tiene una mejor capacidad de pronóstico para el corto plazo (periodo máximo de un año) y para el mediano plazo (de uno hasta máximo 3 años).
- Determinar la validez de la utilización de modelos RNA como una alternativa válida a los modelos ARIMA y VAR para el pronóstico del IPC del Ecuador en el periodo posterior a la dolarización o si por el contrario se obtienen mejores resultados a través de la utilización de estas herramientas de manera complementaria.

5.2.3 Hipótesis

- Los modelos de RNA nos brindaran resultados válidos tras su aplicación para el pronósticos del IPC Ecuador post dolarización.

⁸² Una mayor especificación acerca de las conclusiones rescatadas en el presente documento, pueden ser obtenidas del estudio llevado a cabo por Steven Gonzales denominado *“Neural Networks for Macroeconomics Forecasting: A Complementary Approach to Linear Regression Models”*.

- Las predicciones realizadas por los modelos de RNA resultan ser mejores que aquellos obtenidos a través de los modelos ARIMA y VAR
- La funcionalidad de los modelos de RNA se potencia cuando estos son utilizados como un complemento a los modelos tradicionales.

5.3 Variables y Datos

Tomando en consideración las diferentes corrientes metodológicas analizadas con anterioridad en el presente documento para poder definir los determinantes de la inflación (variables monetarias y de política económica) y en base a los estudios llevados a cabo por Erraez en el año 2004 y por Gachet en el 2007, mismo que fueron publicados en la revista económica del Banco Central del Ecuador, las variables a utilizarse dentro de los diferentes modelos son las siguientes:

VARIABLE DEPENDIENTE.

Inflación: Obtenida directamente desde la página del Banco Central del Ecuador

VARIABLES EXPLICATIVAS.

Tipo de cambio de Colombia: Variable cuyos datos han sido obtenidos directamente del Banco Central de Colombia, dicha variable es tomada en consideración ya que este es uno de los países fronterizos del Ecuador, por lo que el tipo de cambio de esta economía podría llegar a repercutir en la inflación de la economía ecuatoriana.

Tipo de cambio de Perú: Variable cuyos datos han sido obtenidos directamente del Banco Central de Perú, esta variable fue tomada en consideración por el mismo motivo que se consideró tomar en cuenta el tipo de cambio con Colombia.

Tipo de cambio Euro-Dólar: Variable cuyos datos se han obtenido del portal Investing.com, especializado en finanzas y cuyos datos resultan tener una gran fiabilidad, esta variable ha sido tomada en consideración ya que Europa es uno de los principales socios comerciales de Ecuador.

Primary commodity price: Para poder utilizar esta variable se obtuvo un promedio de dos indicadores brindados por la página oficial del Fondo Monetario Internacional (FMI)⁸³, los indicadores tomados en consideración son el Food Price Index (PFOOD) y el Beverage Price Index (PBEVE), se decidió obtener este promedio debido a que los 10 principales productos de exportación de la economía Ecuatoriana son tomados en cuenta para la construcción de estos dos indicadores, por lo que tomar en consideración solo uno de ellos podría generar una pérdida importante de información sobre el efecto que podrían tener el precio de esos bienes para determinar la inflación nacional.

Índice precio del petróleo (POILAPSP): Variable obtenida directamente de la página del FMI, dicho indicador es construido como un promedio simple de los 3 precios de petróleo más importantes (Brent, West Texas y Dubai Fateh), dicha variable es tomada en consideración debido a que este es el principal producto de exportación de la

⁸³ Este indicador puede ser obtenido directamente de la página del Fondo Monetario Internacional, disponible en la siguiente dirección electrónica:
<http://www.imf.org/external/np/res/commod/index.aspx>

economía ecuatoriana, por lo que se espera que variaciones en el precio de este producto tengan repercusión dentro del desempeño económico de la nación y por lo tanto del IPC.

Precio de la tonelada de banano: Variable obtenida directamente de la página del FMI, dicho precio es considerado en base al precio de la tonelada de banano de Centro América y del Ecuador, se ha decidido tomar en cuenta esta variable ya que el banano es el segundo producto no petrolero más importante para el Ecuador, por lo que fluctuaciones en su precio podrían tener efectos dentro del IPC nacional.

Índice de precios de insumos industriales (PINDU): Variable obtenida directamente de la página del FMI, esta variable es un índice construido en base a diferentes insumos utilizados para la producción industrial, tales como las materias primas agrícolas, el precio de los metales y el precio de los materiales de producción.

Ingresos petroleros: Variable obtenida de la página del Banco Central del Ecuador (BCE), esta variable recoge el total de los ingresos percibidos por el estado por concepto de ingresos de la exportación petrolera y por refinados del petróleo.

Ingresos no petroleros: Variable obtenida de la página del BCE, esta variable recoge el total de los ingresos percibidos por el estado que no están relacionados con el petróleo, tanto esta variable como la variable correspondiente a los ingresos petroleros han sido tomados en consideración debido a que se espera que una mayor cantidad de ingresos signifique una mayor consumo de bienes y servicios por parte del estado, incidiendo de esta manera en el IPC de la nación.

Gastos corrientes gobierno central: Variable obtenida de la página del BCE, esta variable recoge gastos del gobierno tales como intereses sueldos o transferencias, por lo que se puede esperar que una variación en esta variable tenga la capacidad de afectar el IPC, ya sea por la variación en los salarios del sector público o por la cantidad de bienes y servicios que este consuma.

Gastos de capital del gobierno central: Variable obtenida de la página del BCE, variable que recoge todos los gastos que lleve a cabo el estado por concepto de la formación bruta de capital fijo, esta variable se encuentra muy ligada al desempeño del sector de la construcción, el mismo que tiene la capacidad de dinamizar una economía, por lo que se espera que el cambio en esta variable se relacione con cambios en la inflación.

Salario básico unificado: Variable obtenida del ministerio de trabajo, esta variable nos permite en cierta medida medir la capacidad adquisitiva de la población y con ella la demanda de bienes y servicios de la misma.

Procesos democráticos: Variable categórica obtenida a través de la revisión histórica de los procesos electorales democráticos que se han llevado a cabo en el Ecuador desde el año 2000, la misma que toma el valor de 1 si en ese mes ha existido un proceso democrático y 0 si no. Esta variable es elegida debido a que se ha identificado a la misma como un factor determinante de la inflación según el trabajo llevado a cabo por

la SEMPLADES denominado “*Políticas públicas para el control de la inflación*” (SEMPLADES, 2008).

Meses de inflación estacionaria: Es una variable categórica construida a partir del análisis histórico de los meses donde se presenta una mayor inflación debido a festividades o desarrollo de actividades como el inicio del año escolar, dicha variable adquiere el valor de 1 si corresponde al mes que presenta incrementos de inflación estacionarios y 0 para otros meses.

Tasa de interés activa referencial: Variable obtenida de los boletines mensuales del BCE, la variable específica tomada en consideración es la tasa de interés corporativa, esta variable es tomada en consideración debido a que la tasa de interés al cual el sector corporativo nacional puede conseguir un crédito, podría influir directamente en el precio de los bienes y servicios ofrecidos por este.

Índice de actividad económica coyuntural (IDEAC): Variable obtenida del BCE, la cual muestra el desenvolvimiento del sector productivo nacional, variable ha sido tomada en cuenta ya que es natural que ante un mejor desarrollo de la actividad económica, se observe un incremento sostenible y saludable de la inflación en el país.

Crédito del sector privado: Variable obtenida del BCE, nos muestra cual ha sido el monto que las instituciones financieras privadas han inyectado a la economía ecuatoriana a través de préstamos, esta variable es utilizada debido a que los créditos son una de las maneras que tienen los agentes económicos para poder realizar su consumo, por lo que un incremento de estos puede repercutir positivamente en el incremento de la inflación.

5.4 Descripción de la Metodología

Dado que el artículo planteado consiste en la exploración de cuál es el método que mejores resultados ofrece tanto para modelación como para pronóstico del IPC del Ecuador en el periodo posterior a la dolarización, se desarrollaran 3 métodos, el método ARIMA, el método VAR y el método de RNA.

5.4.1 Modelo ARIMA

Método que nos permite modelar una variable como una función lineal de la variable rezagada y errores debidos al azar a más de ciertos componentes cíclicos o estacionarios, dichos modelos cuentan con 3 componentes:

- Componente auto regresivo (AR)
- Componente integrado (I)
- Componente de media móvil (MA)

La metodología a utilizarse para la modelación es una de las más utilizadas para llevar a cabo la modelación de series de tiempo a través de los modelos ARIMA, dicha metodología es la propuesta por Box y Jenkins la cual consta de las siguientes etapas:

1. Identificación y selección del modelo
2. Estimación del correspondiente modelo

3. Validación del modelo

Para llevar a cabo la medición de la eficiencia de este modelo se procede a separar valores de la serie que nos servirán como valores de contraste y a estimar dichos valores a través del modelo validado, una vez realizado esto se comprueba que tan bien se ajustan los valores predichos con el modelo con los valores reales de la serie.

5.4.2 Modelo VAR

Es un modelo de ecuaciones simultáneas donde todas las variables son tratadas sin una distinción a priori entre variables endógenas y exógenas, este modelo está formado por un sistema de ecuaciones de forma reducida donde las variables contemporáneas no son las variables explicativas, sino sus rezagos, este modelo esta expresado de la siguiente manera:

$$\begin{aligned}X_{1t} &= \alpha + \alpha_1 X_{1(t-i)} + \alpha_2 X_{2(t-i)} + \dots + \alpha_n X_{n(t-i)} + e_{1t} \\X_{2t} &= \beta + \beta_1 X_{1(t-i)} + \beta_2 X_{2(t-i)} + \dots + \beta_n X_{n(t-i)} + e_{2t} \\&\vdots \\X_{nt} &= \gamma + \gamma_1 X_{1(t-i)} + \gamma_2 X_{2(t-i)} + \dots + \gamma_n X_{n(t-i)} + e_{nt}\end{aligned}$$

Donde:

X_{jt} : Corresponde a las variables que forman parte del modelo.

α, β, γ : Corresponde a cada uno de los parámetros estimados para cada ecuación.

i : Número de rezagos necesarios utilizados en el modelo.

e_{nt} : Terminio de error.

O de forma simplificada el modelo puede ser escrito como:

$$X_t = \beta + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \dots + \beta_n X_{t-n} + e_t$$

Donde β representa el vector de parámetros estimados, X representa el vector de variables, e_t representa el vector de términos de error y $t - n$ representa el número de rezagos utilizado para cada vector de variables.

La estimación de este modelo se realiza a través de MCO aplicado a cada ecuación por separado, lo cual nos generara unos estimadores consistentes siempre y cuando los errores sean ruido blanco. De esta manera, primero se realiza la estimación de la ecuación

$X_{1t} = \alpha + \alpha_1 X_{1(t-i)} + \alpha_2 X_{2(t-i)} + \dots + \alpha_n X_{n(t-i)} + e_{1t}$ a través de MCO y posteriormente se realiza la estimación de la ecuación $X_{2t} = \beta + \beta_1 X_{1(t-i)} + \beta_2 X_{2(t-i)} + \dots + \beta_n X_{n(t-i)} + e_{2t}$ también por MCO, se estimarán tantos modelo como variables tenga el sistema. La utilización de dicha metodología se justifica ya que a diferencia de otros modelos de ecuaciones simultaneas, en los modelos VAR no existen problemas generados por la mala especificación de variables exógenas.

Sin embargo es importante destacar que los pronósticos de los modelos VAR resultan en muchas ocasiones ser confiables solamente en el corto plazo (estos modelo pueden demostrar cierta debilidad en los pronósticos de largo plazo). Para llevar a cabo el pronóstico se suele utilizar el pronóstico dinámico (Londoño, 2005), el cual consiste en llevar a cabo la estimación del modelo con los datos dados y a partir de esta estimación pronosticar varios valores fuera de muestra, cada uno de los pronósticos del periodo $t+1$ se basan en el pronóstico del periodo t , es por este motivo que mientras mayor sea el plazo de tiempo sobre el que se quiere realizar el pronóstico, menos confiable resulta este.

Para llevar a cabo el pronóstico óptimo de las serie fuera de muestra requerimos conocer nuestro vector de parámetros estimados y nuestra matriz de varianzas covarianzas, los mismos que pueden ser estimados a partir de los datos disponibles hasta el periodo n (en el caso del estudio propuesto se utilizaran los datos mensuales hasta el año 2016) y a partir de este año realizar los pronósticos para los periodos posteriores.

Es importante mencionar que con la finalidad de realizar pronósticos que resulten confiables tanto en el corto como en el mediano plazo, se debe asegurar que las relaciones obtenidas a través de modelo VAR no resultan ser relaciones espurias. Para esto se procederá a probar la cointegración de las series a través del análisis de cointegración de Engle-Granger, lo cual nos brindaría la confianza de que el modelo resulta adecuado para llevar a cabo pronósticos para periodos más largos de tiempo.

De esta manera el pronóstico de nuestra serie para el periodo $t+1$ será el valor esperado de X_{t+1} dado los valores de X_1, X_2, \dots, X_t . Este resultado puede ser expresado de la siguiente manera:

$$\hat{X}_{t+1} = E(X_{t+1} | X_1, X_2, \dots, X_t) = \beta X_t$$

De la misma manera, para llevar a cabo el pronóstico del segundo periodo se procede de la misma manera, teniendo el siguiente resultado:

$$\hat{X}_{t+2} = E(X_{t+2} | X_1, X_2, \dots, X_t) = \beta^2 X_t$$

Dicho resultado puede ser generalizado para los pronósticos que periodos posteriores al segundo, de la siguiente manera:

$$\hat{X}_{t+i} = E(X_{t+i} | X_1, X_2, \dots, X_t) = \beta^i X_t$$

5.4.3 Modelo de RNA

Los modelos de RNA son modelos computacionales cuyo funcionamiento está basado en el comportamiento simplificado del cerebro humano y la conexión existente entre las neuronas biológicas que lo conforman. Al tener un funcionamiento análogo al cerebro humano, el funcionamiento de las RNA replica la sinapsis neuronal existente en el mismo y a través de este proceso lleva a cabo las funciones para los cuales fue diseñado, ya sean estas procesos de análisis de datos, inteligencia artificial, pronósticos de series temporales, identificación de patrones, etc.

Es precisamente la similitud que tiene el funcionamiento de los modelos de RNA con el funcionamiento del cerebro humano lo que le permite a estas herramientas obtener ciertas ventajas tales como:

- Mantienen un aprendizaje adaptativo lo cual implica que estos modelos puedan aprender a través de la experimentación, denominado etapa de aprendizaje.
- Tiene la capacidad de auto organizarse en base a los resultados obtenidos en la etapa de aprendizaje.
- Tienen una alta tolerancia a los fallos que se puedan presentar dentro de la estructura del modelo de RNA.
- Su aplicación resulta menos restrictiva que otros tipos de modelos utilizados para el pronóstico debido a que no es necesario la adopción de supuestos a priori

Debido a la capacidad de aprendizaje adaptativo de estos modelos, no es necesario asumir a parámetros fijos ni tampoco es necesario la presunción de la manera en la que los datos utilizados se encuentran distribuidos, esto permite que el investigador tenga que enfocar su atención en la correcta selección de la arquitectura del modelo de RNA a utilizarse.

Los elementos de los cuales se encuentran compuestos una RNA son: la función de entrada, los pesos sinápticos, la función de activación y la función de salida.

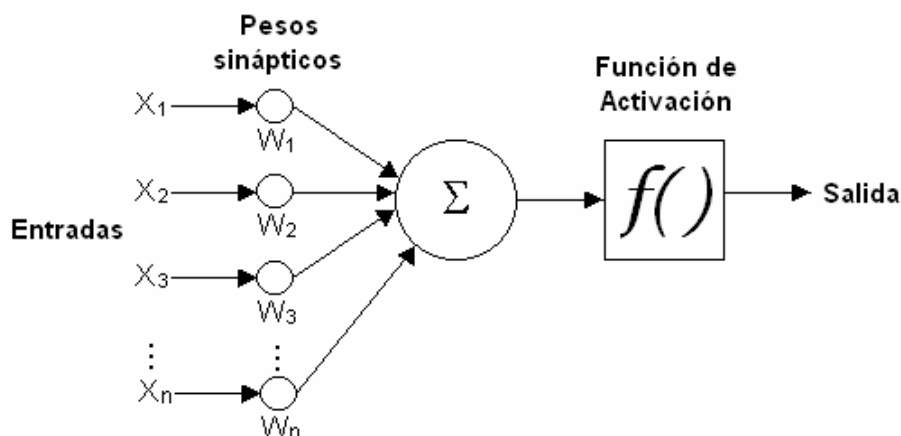


Figura 1: Funcionamiento del modelo de RNA

Fuente: José Castro, 2006, Fundamentos para la implementación de una red neuronal perceptron multicapa, San Carlos, Guatemala, Facultad de Ingeniería, Universidad de San Carlos de Guatemala.

- **Función de entrada:** Es la capa en la que se ingresan los datos externos de la red, mismo que son tratados por las diferentes neuronas como una sola entrada global a través de una función de entrada de la siguiente forma:

$$\text{funcion de entrada} = (X_1 W_1) * (X_2 W_2) * \dots * (X_i W_i)$$

Donde * representa el operador a utilizarse para llevar a cabo la relación entre las diferentes variables de la función de entrada, dichos operadores pueden ser: sumatoria (Σ), pitatoria (\prod), mínimo valor o máximo valor.

A partir de esta relación se consigue obtener el nivel de influencia de cada una de las entradas dentro de los valores obtenidos en la capa de salida y realizar el

ajuste de este grado de influencia a través de la modificación de los respectivos pesos sinápticos relacionados a cada entrada.

- **Pesos Sinápticos:** Está compuesto por valores numéricos fijos asignados por el programa en función de la relación que mantenga cada una de las variables de entrada con la función de activación identificada en el modelo, estos pesos pueden ser valores tanto positivos como negativos, la importancia o peso dentro del modelo dependerá de los resultados que se vaya arrojando en la etapa de aprendizaje de la red neuronal.
- **Función de Activación:** Las neuronas artificiales utilizadas dentro del modelo pueden o no estar en un estado de activación dependiendo de su valor de entrada menos el valor umbral, mismo que determinara si la neurona se encuentra o no activa y en función de este estado se transmite información a las demás neuronas que comprenden el modelo. Dentro de esta capa se elige la relación existente entre las variables y sus respectivos pesos sinápticos de tal manera que nos brinde un resultado en la salida de la RNA, para el caso planteado del pronóstico del IPC se podría expresar de la siguiente manera:

$$IPC_i = f \left[\sum_{i=1}^n (X_i \times W_i) \right]$$

Donde: $f(\cdot)$: Función de activación

X_i : Variables explicativas del IPC

W_i : Pesos sinápticos asignados a cada variable

- Cabe recalcar que las funciones de activación para el caso de la RNA perceptron multicapa utilizadas comúnmente son cinco: función limitador fuerte, función limitador fuerte simétrico, función lineal, función sigmoidea logarítmica y función sigmoidea logarítmica hiperbólica (Garcia, 2006), sin embargo, no existe un procedimiento establecido que nos permita elegir entre una función u otra, sino que esta función depende de la naturaleza de los datos y de la experiencia del investigador, para el estudio propuesto solo se tomaran en consideración las tres últimas debido a la naturaleza de los datos.
- **Función Lineal:** es la forma funcional más simple que se puede asumir, la misma que se representa como $f(x) = x$ y por lo tanto no se asigna ningún tipo de límite en la capa de salida. Para el caso del estudio sobre el IPC del Ecuador se prevé que la asignación de esta forma funcional como función de activación, pese a no ser errónea es la que nos puede arrojar resultado más imprecisos ya que la inflación en el Ecuador pos dolarización ha sido muy estable y en el caso de los últimos años ha permanecido en valores muy cercanos a cero, por lo que se espera que las dos funciones de activación restantes brinden resultados en la capa de salida mucho más ajustados a la realidad.
- **Función Sigmoidea Logarítmica:** es una función continua la cual se encuentra comprendida entre valores de cero y uno, esta característica hace que la función brinde como valores de activación a aquellos valores comprendidos entre este

rango, la forma funcional de la función sigmoidea logarítmica es $f(x) = \frac{1}{1+e^{-x}}$.

Se espera que los resultados obtenidos en la capa de salida a través de la utilización de esta función de activación sean mejores que a través de la utilización de la función lineal debido a la naturaleza propia de los datos con los que se está trabajando.

- **Función Sigmoidea Tangente Hiperbólica:** Al igual que la función sigmoidea logarítmica, esta también es una función continua, por lo que es diferenciable en todo su dominio, sin embargo a diferencia de esta, la función sigmoidea tangente hiperbólica tiene como valores de activación aquellos comprendidos entre uno y menos uno, su forma funcional está dada por $f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$. Dado un rápido análisis de los datos con los que se trabajan, se espera que los resultados obtenidos tras la aplicación de esta función de activación, sean menos precisos que los que se obtienen con la aplicación de la función sigmoidea logarítmica, pero más precisos que aquellos obtenidos con la aplicación de la función lineal.
- **Función de Salida:** Es el ultimo componente que tiene una neurona artificial, esta función es la que proporciona el valor de salida de la neurona artificial, mismo que puede ser el valor de entrada para otras neuronas o el valor resultante del modelo de la red neuronal, este valor esta dado en base a la función de activación, la misma que de encontrarse por debajo del valor umbral no activara la neurona y por lo tanto no se enviará ningún valor a la función de salida, en caso de que la función de activación brinde resultados superiores al valor del umbral entonces se enviará un impulso hacia la función de salida, lo que arrojará un valor de salida de la neurona comprendido entre cero (o -1 si la función de activación fuera la sigmoidea hiperbólica) y 1.

La eficiencia de este tipo de modelos depende del número de capas ocultas que se creen dentro de la red neuronal y el número y peso de las interconexiones existentes entre las neuronas que conforman cada capa, ya que si el número de capas y de neuronas es el adecuado los resultados obtenidos tendrán muy buenas propiedades, aunque desafortunadamente no existe una forma técnica de identificar el numero correcto de capas y neuronas necesarias, sino que esto más bien depende de la experiencia del investigador.

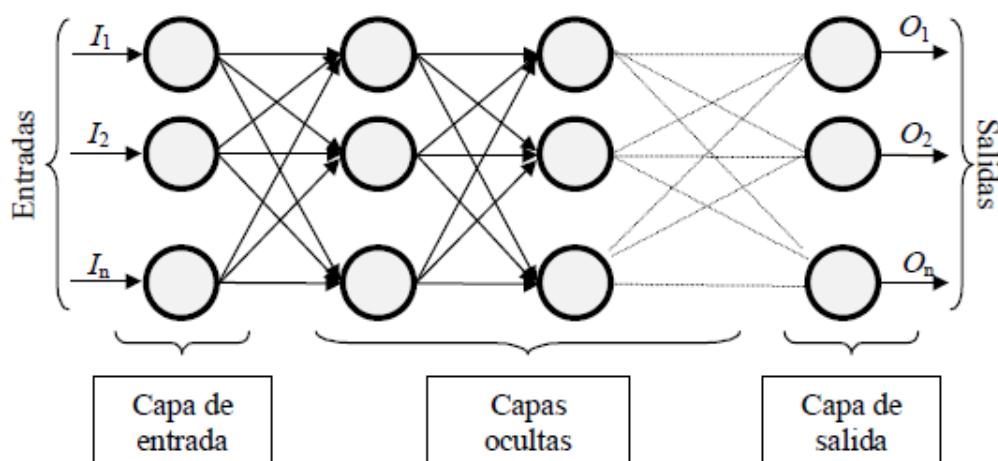


Figura 4: Estructura del modelo de RNA

Fuente: Carlos R., Marta B., 2001, Redes Neuronales: Conceptos Básicos y Aplicaciones, Rosario, Argentina, Departamento de Ingeniería Química, Universidad Tecnológica Nacional.

Una de las etapas más importantes para el correcto funcionamiento de los modelo RNA es la etapa de aprendizaje, en esta etapa los pesos sinápticos asignados se van modificando de tal manera que los erros obtenidos en la capa de salida de la RNA sean lo más cercanos posibles a los valores reales, es decir que los datos obtenidos a través de la RNA son validados en el proceso de aprendizaje lo que permite que no se requiera de un algoritmo que permita llevar a cabo las funciones requeridas, sino que este se va generando a través de la modificación de los pesos mediante el aprendizaje, este procedimiento se repite hasta que la minimización de los errores entre cada prueba deje de variar. Dentro del proceso de aprendizaje de la RNA se pueden identificar dos componentes los cuales son el paradigma de aprendizaje y el algoritmo de aprendizaje:

Paradigma de Aprendizaje

Dado que las nuevas aplicaciones de RNA llevan a cabo los procesos de minimización del error entre los valores predichos y los valores esperados a través de la modificación de los pesos sinápticos (para las redes que mantienen los pesos sinápticos fijos no existe la etapa de entrenamiento) la elección de un correcto paradigma de aprendizaje resulta un aspecto muy relevante en la capacidad de la RNA para llevar a cabo la tarea para la cual ha sido creada, estos paradigmas pueden ser el aprendizaje supervisado y el aprendizaje no supervisado.

- **Aprendizaje Supervisado:** este proceso de aprendizaje es el más utilizado dentro de los modelo RNA, especialmente en los modelo multicapa, en dicho proceso existen dos tipos de datos; los primeros que permiten a la red neuronal generar las conexiones entre las neuronas y asignar los correspondientes pesos sinápticos, es decir construir el modelo de redes neuronales a utilizarse. Una vez se haya obtenido está RNA se procede a validar la misma a través de la inserción de valores de entrada y las correspondientes salidas de la misma, estos valores

de salida son contrastados con el segundo grupo de datos que son asignados como valores de control.

En base al contraste entre los valores predichos y los valores deseados se encuentra una medida del error obtenido a través del uso de la RNA, se procede a modificar los pesos sinápticos asignados dentro del modelo y se lleva a cabo el proceso señalado anteriormente, este proceso se repite hasta que el error obtenido entre los valores predichos y los valores deseados sea el mínimo o se haya llegado a un punto en el cual no exista diferencia entre la simulación del periodo n y la simulación del periodo $n+1$.

- **Aprendizaje no Supervisado:** a diferencia del entrenamiento supervisado, en este caso no existen datos que le indiquen a la RNA los valores de salida con los cuales se deben contrastar los resultados sino que el modelo solo cuenta con valores de entrada, es decir que el modelo no cuenta con información externa que le permita saber si las salidas de la red son los valores deseados.

Este tipo de modelo lo que hacen es ir estableciendo una serie de conexiones, correlaciones, similitudes, categorías, distribuciones, entre otras, que le permitan encontrar patrones a través de los cuales ir modificando su estructura de manera tal que esta se vaya comportando de manera similar al modelo que se tiene con los valores de entrada dado.

Algoritmo de Aprendizaje

Los algoritmos de aprendizaje son las partes de la estructura del modelo de redes neuronales el cual se encarga de modificar los pesos asignados a cada valor de entrada de manera tal que los valores obtenidos en la función de salida se vayan aproximando cada vez más a los valores esperados, disminuyendo de esta manera el error del modelo a su mínimo.

- **Algoritmo Backpropagation:** Son algoritmos ampliamente utilizados en las redes multicapa de tipo perceptron. Presentan una gran fortaleza ya que dichos algoritmos realizan un numero grande de iteraciones de tal manera que en cada una de ellas se tienda a reducir el error, por lo general este algoritmo lleva a cabo una metodología conocida como el descenso del gradiente⁸⁴ del error cuadrático medio el cual modifica los pesos sinápticos en cada proceso iterativo (Areli, 2012), de la siguiente manera:

$$W_{k+1} = W_k + \mu(-\nabla_k)$$

Con:

$$\nabla_k = \frac{\partial e_k^2}{\partial W_k}$$

- **Algoritmo Hebbiano:** lo que busca este algoritmo es extraer generalidades de los datos utilizados dentro de la función de entrada. Este algoritmo compara los niveles de activación (estado activo o inactivo) de cada una de las neuronas dentro del modelo e incrementa la ponderación total de los mismos en base a

⁸⁴ Se debe aclarar que en este caso el gradiente es el vector de las derivadas parciales de la función de los errores con respecto a cada uno de los pesos sinápticos asignados a cada valor de entrada.

estos, es decir que si dos neuronas se encuentran activas, el peso que tiene la conexión entre las mismas se verá incrementada.

- **Algoritmo Competitivo:** este algoritmo procede a agrupar cada una de las entradas en clusters y a partir de estos generar una agrupación de cada una de los nuevos patrones que se vayan presentando dentro del modelo, es decir que si se determina un nuevo patrón, este algoritmo lo clasificara dentro de un cluster existente de ser posible o creara otro en caso de que este no pueda ser incluido en ningún cluster previamente establecido (Carlos Ruiz, 2001), esto permite que conforme se vayan incluyendo nuevos patrones dentro de cada cluster, este vaya adquiriendo una mayor jerarquía dentro del modelo.

Para determinar los pesos que se utilizaran como valores semillas es común establecer un rango máximo y mínimo dentro del cual los pesos para cada valor de entrada se van a encontrar, lo más común es que dichos valores sean valores pequeños, entre un rango de -0.5 y +0.5, aunque en la fase de entrenamiento los pesos no tiene por qué verse restringidos a este intervalo. El proceso de aprendizaje se detendrá una vez la modificación de los pesos sinápticos entre las neuronas arroje como resultado un error cuadrático medio mínimo (cercano a cero) o que la diferencia entre el RMS de la iteración k y la iteración $k-1$ sea casi inexistente, una vez se haya obtenido esta RMS deseada los pesos sinápticos son establecidos como los pesos finales del modelo.

Tratamiento de los datos de entrada

Previo a la utilización de los datos para la generación del modelo de RNA, estos deben ser tratados de tal manera que se permita distinguir las características que permitan identificar los valores de activación de las neuronas relacionados a los mismos, para llevar a cabo este tratamiento se deben diferenciar entre datos que presentan atributos continuos y discretos .

- **Atributos continuos:** cuando los datos con los que se trabajan corresponden a datos de series de tiempo (como los utilizados en el presente estudio) es necesario que estos se codifiquen dentro de un intervalo comprendido entre valores de 0 y 1 correspondientes al valor máximo y al valor mínimo de la serie, esto se logra a través de una función lineal, la misma que nos permite también considerar los valores numéricos del modelo que puedan ubicarse tanto por encima como por debajo de los máximos y mínimos observados hasta ese momento en la serie.
- **Atributos discretos:** para el caso de las variables que representen atributos de la población, se deben convertir estas variables en variables dummy correspondientes a ese atributo y asignar una neurona para cada uno de estas variables. Otra alternativa aparece cuando se trata de una variable discreta que presenta jerarquía, en este caso se podría dividir al valor de 1 (los valores permitidos en las entradas solo está entre un rango de 0 y 1) entre el número de categorías que tenga la variable y en función de esto generar una nueva variable que represente a estos atributos, por ejemplo si se tiene una variable cuyo

atributo sea regular, bueno, muy bueno y excelente, la variable transformada comprendida entre 0 y 1 sería: de 0 a 0.25 regular, 0.25 a 0.5 bueno, de 0.5 a 0.75 muy bueno y de 0.75 a 1 excelente.

5.4.4 Criterios de evaluación de los modelos

Para poder responder de manera objetiva a las preguntas de investigación planteados en el presente documento, lo primero que se llevara a cabo es la estimación de los diferentes modelos propuestos (ARIMA, VAR, RNA) utilizando para ellos los datos disponibles desde el periodo comprendido entre los años 2000-2017, posteriormente se llevara a cabo un pronóstico a través de cada modelo sobre la serie correspondiente a los últimos 12 o 24 meses correspondientes al corto y el mediano plazo (2016-2017 o 2015-2017).

Para encontrar respuesta a la primera pregunta de investigación (validez de los modelos RNA para el pronóstico del IPC) se puede recurrir a diversas técnicas de evaluación como una medida de bondad de ajuste medido por el R^2 (Arturo Morales, 2008) o técnicas estadísticas como el criterio de Akaike, criterio de Schwarz o el criterio de información Bayesiano (Aristizabal, 2006), mismos que nos resultan útiles para determinar que tan bien se ajustan los modelos a los datos con los que se trabaja dentro de muestra.

Con la finalidad de responder a la segunda pregunta de investigación (¿Cuál es el modelo que mejores pronósticos ofrecen?), se llevara a cabo el pronóstico de los últimos 12 y 24 periodos (correspondientes al corto y mediano plazo) a través de cada modelo, con la finalidad de contrastar posteriormente los resultados. Para realizar este análisis se recurrirá, en primer lugar, a la construcción de los gráficos correspondientes a los valores pronosticados por cada modelo contrastado con la serie original, posteriormente se llevara a cabo un análisis del error cuadrático medio, de la misma manera se plantea la posibilidad de utilizar otros estadísticos como la U de Theil para llevar a cabo este análisis.

Finalmente para responder la tercera pregunta de investigación (las RNA trabajan de mejor manera como una alternativa o como un complemento con los modelos tradicionales) se debe realizar primero una breve explicación.

Como se mencionó anteriormente, para llevar a cabo la selección del modelo RNA más eficiente no existen parámetros pre establecidos, por lo que realizar el análisis del funcionamiento de cada modelo de RNA utilizando diferentes variables puede ser un trabajo exhaustivo e innecesario. En base a esta premisa, en el presente estudio se pretende construir un modelo RNA con todas las variables disponibles y utilizar todas las pruebas mencionadas anteriormente (tanto de validez como de capacidad de pronóstico) y posteriormente construir un modelo RNA tomando en cuenta solo las variables explicativas que resulten ser significativas en el modelo VAR, así como los rezagos y especificaciones obtenidas tanto del modelo ARIMA como del modelo VAR, posteriormente se realizarán las mismas pruebas a este modelo y se llevará a cabo el contraste de resultados.

6. Contenido Mínimo

- Resumen
- Abstract
- Introducción
- Marco teórico
- Revisión de la Literatura
- Metodología
- Resultados
- Conclusiones y recomendaciones
- Anexos
- Bibliografía.

7. Cronograma de Actividades

Actividades	Tiempo
Revisión de la literatura	4 semanas
Revisión y ajuste de los datos a utilizarse	1 semana
Corrida de modelo ARIMA, optimización del modelo, pruebas de ajuste, pronóstico	1 semana
Corrida de modelo VAR, optimización del modelo, pruebas de ajuste, pronóstico	1 semana
Corrida de modelo RNA, optimización del modelo, pruebas de ajuste, pronóstico	1 semana
Contraste de resultados obtenidos (ajuste de los modelos y pronósticos) y ajustes finales a los modelos	4 semanas
Redacción del artículo	4 semanas

8. Bibliografía

Alberto Aragón, J. M. (17 de Octubre de 2014). Aplicaciones de Redes Neuronales en Economía . Burgos, España : Universidad de Burgos .

Andrew Ang, G. B. (lunes de Marzo de 2006). Do Macro Variables, Asset Markets, or Surveys ? . New York , Estados Unidos : Federal Reserve Board of Governors, Division of Monetary Affairs.

Arelí, J. (abril de 2012). CUCEI. Obtenido de http://www.cucei.udg.mx/sites/default/files/pdf/toral_barrera_jamie_areli.pdf

Argandoña, A. (Julio de 1990). El pensamiento economico de Milton Friedman . Barcelona, España: IESE Bussines School .

Aristizabal, M. (Noviembre de 2006). Evaluacion asimetrica de una red neuronal artificial: aplicacion al caso de la infalcion en colombia . Antioquia, Colombia : Universidad de Antioquia .

Arturo Morales, I. C. (11 de Noviembre de 2008). Pronostico mediante redes neuronales artificiales y modelos ARIMA: el caso de los cetes en Mexico . Mexico , Mexico : Facultad de contaduria y administracion UNAM.

- Bruce Baker, C. R. (martes de Febrero de 1999). A comparison of conventional linear regression methods and neural networks for forecasting educational spending. New York , Estados Unidos : Pergamom .
- Carlos Ruiz, M. B. (Marzo de 2001). Redes Neuronales: Conceptos Básicos y Aplicaciones . Rosario , Argentina : Departamento de Ingeniería Química .
- Elena Grubisic, M. M. (Agosto de 2000). Modelos de predicción de Agregados Monetarios. Buenos Aires , Argentina : Repositorio Banco Central .
- Erraez, J. P. (Septiembre de 2005). El proceso inflacionario en el Ecuador: un análisis de sus determinantes con modelos Arima y Vectores autorregresivos. Quito , Ecuador : Banco Central del Ecuador Cuestiones Economicas.
- Erraez, J. P. (2007). *El proceso inflacionario en el Ecuador: Un analisis de sus determinantes con modelos Arima y Vectores autoregresivos* . Quito : Banco Central .
- Francesco Bianchi, A. C. (Jueves de Julio de 2014). Globalization and Inflation: Structural Evidence from a Time Varying VAR approach . Estados Unidos : Elsevier.
- Garcia, J. F. (Noviembre de 2006). FUNDAMENTOS PARA LA IMPLEMENTACIÓN DE RED NEURONAL. Guatemala , Guatemala : Universidad San Carlos .
- George Box, G. J. (1976). *Time Series Analysis forecasting and control* . California : Holden-day.
- Gonzalez, S. (2007). Neural Networks for Macroeconomic Forecasting: A Complementary Approach to Linear Regression Models. Ottawa, Canada : Government of Canada Publications.
- Gregorio, J. d. (2007). *Macroeconomía Teoría y Políticas*. Santiago : Pearson Educacion .
- Instituto Nacional de Estadísticas y Censos . (2015). *Metodología de índice de precios al consumidor (IPC) Base anual: 2014=100*. Quito : INEC.
- Ivan Gachet, D. M. (Febrero de 2008). Determinantes de la Inflación en una Economía Dolarizada: El Caso Ecuatoriano . Quito , Ecuador : Banco Central del Ecuador .
- Keynes, J. M. (Octubre de 1924). A track on monetary reform . Londres, Reino Unido : McMillan and Co.
- Londoño, W. (24 de Noviembre de 2005). Modelo de ecuaciones multiples, modelo VAR y cointegracion . Medellin, Colombia : Universidad EAFIT.
- Nakamura, E. (16 de Septiembre de 2004). Inflation forecasting using a neural network. Massachusetts , Cambridge, Estados Unidos : ELSEVIER.
- OIT / FMI / OCDE / Oficina Estadística de las Comunidades Europeas / Organización de las Naciones Unidas / Banco Mundial. (2006). *Manual del índice de precios al consumidor: Teoría y práctica*. Washington : Departamento de Tecnología y Servicios Generales.
- Pablo Garcia, R. V. (Diciembre de 2003). *Cuaderno de economía version online*. Obtenido de Scielo.com: http://www.scielo.cl/scielo.php?pid=S0717-68212003012100042&script=sci_arttext&lng=pt



Paul McNelis, P. M. (Abril de 2004). FORECASTING INFLATION WITH THICK MODELS AND NEURAL NETWORKS . Frankfurt , Alemania: Banco Central Europeo .

SEMPLADES. (Octubre de 2008). Politicas Publicas Para el Control de la Inflacion . Quito , Ecuador : Semplades.

Sims, C. A. (1980). Microeconomics and Reality . *Econometrica* , 1-48.